# *Papers in Austronesian subgrouping and dialectology*



Bowden, J. and Himmelmann, N. editors. *Papers in Austronesian subgrouping and dialectology*.
**PL-563**, viii + 169 pages. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.cover*

# Pacific Linguistics 563

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

# Papers in Austronesian subgrouping and dialectology

John Bowden and Nikolaus Himmelmann, eds

Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia



First published 2004

National Library of Australia Cataloguing-in-Publication entry:

John Bowden and Nikolaus Himmelmann, eds
Papers in Austronesian subgrouping and dialectology

Bibliography.
ISBN 0 85883 477 2

1. Austronesian languages – Variation. 2. Austronesian languages – Dialectology
I. Bowden, John (Frederick John), 1958– . II. Himmelmann, Nikolaus, 1959– .
III. The Australian National University. Research School of Pacific and Asian Studies.
IV. Title. (Pacific Linguistics 563).

499.57

Copyedited by Basil Wilson
Typeset by Jeanette Coombes
Cover design by Cirils Printers
Printed Cirils Printers, Fyshwick, Canberra

# *Table of contents*

# *Notes on contributors*

**John Bowden** is a research fellow in linguistics at the Research School of Pacific and Asian Studies at The Australian National University. He is interested in linguistic typology in general, as well as the typology of Austronesian and Papuan languages in eastern Indonesia and East Timor.

**Terry Crowley** teaches at the University of Waikato (New Zealand). In the 1970s he researched Australian languages but has been working on Oceanic languages since then. He has published grammars of Bandjalang and Uradhi in Australia, and of Paamese and Erromangan in Vanuatu. He is currently working on languages of northern Malakula.

**Mark Donohue** teaches at the National University of Singapore. He has worked and published extensively on the languages of eastern Indonesia and New Guinea, including a reference grammar of Tukang Besi and sketches of Warembori and I'saka, and is currently preparing a full-length grammar of Skou.

**Nikolaus Himmelmann** is Professor of Linguistics in the Department of Linguistics, Ruhr-Universität Bochum, Germany. He has done fieldwork in the Philippines (Tagalog), Sulawesi (Tomin-Tolitoli languages) and East Timor (Waima'a) and published widely on a number of core issues in Austronesian grammar, including the nature of lexical and syntactic categories and voice.

**Jae Jung Song** teaches Linguistics at the University of Otago, New Zealand. He studied at Monash University, Melbourne, where he also gained his PhD. He has contributed to international journals including *Lingua*, *Linguistics* and *Oceanic Linguistics*. He is the author of *Causatives and causation* (Addison Wesley Longman 1996) and *Linguistic typology: morphology and syntax* (Pearson Education 2001).

**Malcolm Ross** is a Professor in the Department of Linguistics in the Research School of Pacific and Asian Studies at The Australian National University and Deputy Director of the Centre for Research on Language Change at ANU. He has published extensively on historical linguistics in the Austronesian family.

**Der-Hwa Victoria Rau** is Professor of Linguistics in the Department of English Language, Literature, and Linguistics at Providence University in Taiwan. Her main research interests are in sociolinguistic variation and applied linguistics. Her publications include *A grammar of Atayal* (1992), *Yami textbook* (with Maa-Neu Dong 2000), and a forthcoming book on Yami.

**René van den Berg** is a linguistics consultant with SIL-International. He has worked and published extensively on the Muna language of southeast Sulawesi, including a reference grammar, two dictionaries and the Indonesian translation of an old Dutch ethnographic source. He is currently based in Papua New Guinea.

Bowden, J. and Himmelmann, N. editors. *Papers in Austronesian subgrouping and dialectology*.
**PL-563**, viii + 169 pages. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.cover*

# 1 *Perspectives on subgrouping and dialectology: an introduction*

JOHN BOWDEN AND NIKOLAUS HIMMELMANN

This volume has had a rather unusual genesis. It was originally conceived as part of the now defunct Series A of Pacific Linguistics, which was once used to publish collections of occasional papers. For a number of reasons, a decision was made to abandon the occasional papers series, and only to publish collections of articles which had some sort of thematic coherence. Of course, there was still the problem of how to deal with a few papers which had already been accepted for publication in the A series, in this case the papers by Donohue and Rau that appear in this book. Fortunately, the two papers that had already been accepted for publication both dealt with the classification of languages and dialects in different ways and it was clear that there was already the making of a volume with the kind of thematic coherence Pacific Linguistics wanted.

There was another complication with this volume as well. The book began with just Nikolaus Himmelmann as sole editor. However, when he left The Australian National University to take up a position in Germany, he no longer had the time to devote to the rest of the task of putting this volume together alone. When John Bowden came to the ANU, he joined Himmelmann as coeditor, and put out a call for papers on Austronesian subgrouping and dialectology that could combine with the ones already received to make a coherent volume.

Although this book was not planned in quite the same way as most collections organised around a theme, we hope that the volume as it has turned out will be a worthwhile addition to studies on dialectology and subgrouping in the Austronesian family. The papers reflect a wide range of approaches to looking at questions of linguistic classification at both macro and micro levels.

Terry Crowley's paper compares how linguists have talked about the classification and naming of languages and dialects with how speakers of Oceanic languages themselves actually name and classify their ways of speaking. He takes issue with Mülhäusler's (1996) claim that languages represent colonial abstractions produced by foreign academics and missionaries and have no relationship to indigenous understanding of what their local linguistic situations are like. He cites a wealth of evidence from the Pacific surveying indigenous classification of speech varieties in order to refute Mülhäusler's claim.

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 1–2.
Canberra: Pacific Linguistics, 2004.



Bowden, J. and Himmelmann, N. "Perspectives on subgrouping and dialectology: an introduction". In Bowden, J. and Himmelmann, N. editors, *Papers in Austronesian subgrouping and dialectology*.
**PL-563:1-2**. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.1*

Mark Donohue's paper constitutes an examination of a long-standing claim about the unity of a Muna Buton group of languages consisting of all the languages spoken in southeast Sulawesi. As with quite a few long-standing claims on Austronesian subgrouping that have been repeated many times, no one has yet offered any hard evidence for the group, although speculation has abounded. Donohue assesses a range of phonological evidence and concludes that not all of the languages of this area can be included together in one subgroup. Nevertheless, it is possible to link many of the languages from this area together into a coherent subgroup. However, the Wolio and Kamaru languages do not belong to this group and the status of others not examined in his paper remains unclear.

Victoria Rau's paper aims to clarify the relationships between Atayalic dialects of Taiwan. In her paper she focuses on the Mstbaun, Palngawan and Inago dialects. A feature of her study is the careful attention paid not just to lexical similarity, but also to sound change and mutual intelligibility. Rau concludes that a family-tree model for the relationships between these groups is not appropriate, but that rather they should be seen as a dialect chain with Mstbaun between Palngawan and Inago.

The major contribution of Malcolm Ross's paper is to map out the major features of an internal subgrouping for the Malayic subgroup of Austronesian, mostly using as evidence shared innovations in bound morphology. The Malayic group has largely resisted earlier efforts at internal classification based on phonological innovations because of widespread lexical borrowing between the languages. He concludes that Old Malay, despite its nomenclature, was not actually a part of the Malayic group, and further proposes that Malayic has two major subgroups: a western group including Salako, Ahe and Belangin, and a larger nuclear Malayic group including all the other Malayic languages.

Jae Jung Song also uses the evidence of bound morphology in Micronesian to assess competing claims about subgrouping within that subgroup of Oceanic. Debate has focused on whether or not a 'flat-tree' model of Micronesian subgrouping should be preferred over a more stratified one. Song concludes that the evidence from the focus system and from the possessive pronouns suggests that the flat-tree model, with little internal branching, is to be preferred.

René van den Berg's paper is a contribution to dialectology in the Muna language. Van den Berg's (1989) grammar of the Muna language focused on the most prestigious and most widely spoken northern dialect of the language spoken in Muna subdistrict. The paper in this volume provides quite detailed description of the main features of the lesser known southern dialect which differ from that of the north. In this sense, the paper here could perhaps be viewed as an important appendix to his (1989) grammar.

## References

Mülhäusler, Peter, 1996, *Linguistic ecology: language change and linguistic imperialism in the Pacific region*. London: Routledge.

van den Berg, René, 1989, *A grammar of the Muna language*. Dordrecht: Foris.

Bowden, J. and Himmelmann, N. "Perspectives on subgrouping and dialectology: an introduction". In Bowden, J. and Himmelmann, N. editors, *Papers in Austronesian subgrouping and dialectology*.
**PL-563:1-2**. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.1*

# 2 *The question of dialect and language in Oceania*

TERRY CROWLEY

## 1 Introduction[1]

Given that many of the 450 or so Oceanic languages — nearly 10% of the world's total number of languages — exhibit varying degrees of internal regional diversity, it is perhaps a little surprising that data from so few of these languages have contributed significantly to the field of dialectology in general, and to Austronesian dialectology in particular. In fact, the only major dialectological studies of any Oceanic language that I am aware of relate to Fijian (Schütz 1972; Geraghty 1983). Published grammars and dictionaries of Oceanic languages for the most part concentrate on just a single regional variety, though often with some specific comments on the major points of phonological, lexical or grammatical features by which other regional varieties differ from the described variety, for example Crowley (1982:8–10) and Crowley (1992:x–xvi) for Paamese, with little attempt to describe variability, the effects of dialect contact and dialect levelling.

Any discussion of Austronesian dialectology must, of course, be predicated on some kind of understanding of what constitutes the difference between a dialect and a language. The issue of whether varieties of speech associated with different geographical areas should be considered as 'dialects of a single language' or as 'different languages' is, of course, typically decided by invoking the criterial notion of mutual intelligibility. Geographically determined speech forms which are not mutually intelligible are said to constitute separate languages, whereas dialects of the same language are generally said to be mutually intelligible.

Since a speaker of Fijian cannot understand anything of what is said when somebody is speaking Maori, we can easily say that Fijian and Maori constitute separate languages. On the other hand, someone who has learnt Maori in the East Cape area of New Zealand can easily understand somebody who learned the language in Northland, despite the existence of some recognisable differences between the two varieties, so the speech patterns of East Cape and Northland constitute two dialects of a single language.

1 I would like to thank John Lynch, John Bowden and Jeff Siegel for helpful comments on a preliminary version of this paper. Thanks also to participants in a seminar on this topic at the University of New England (Armidale, Australia) in September 2000, which resulted in interesting discussion. Final responsibility for all interpretation and observations within this paper rests solely with the author.

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 3–20
Canberra: Pacific Linguistics, 2004.

However, it is also generally acknowledged that sociopolitical factors are often at least equally important in distinguishing between different languages, as reflected in the widely repeated[2] aphorism that 'a language is a dialect with an army and a navy'. Dutch, for example, is about as distinct from the German of Berlin as is the local speech of many parts of northern Germany, yet only Dutch is said to constitute a separate language, because of its association with what has become a separate nation. Afrikaans, on the other hand, could easily have been treated as a dialect of Dutch rather than a separate language if it had been spoken in the Netherlands rather than in South Africa. In fact, until relatively recently, even Afrikaaners did call their language Dutch, though they now refer to it as a separate language.

This paper compares how linguists talk about Oceanic languages and the regional diversity to be found within them with how speakers of these languages themselves talk about the same sorts of issues. The paper seeks to establish whether languages are created out of diversity by speakers of those languages themselves, or whether, as argued by Mühlhäusler (1996), languages represent colonial abstractions produced by foreign academics and missionaries which are not in accord with indigenous understandings of the linguistic situation of the region.

## 2 Revisiting dialect and language in Oceania[3]

For several decades after the appearance of Wurm and Laycock's article on the question of language and dialect in New Guinea in 1961, there was no serious attempt to discuss the dialect/language issue in terms that might be applicable to Oceanic languages.[4] Wurm and Laycock (1961:137) concluded that 'the ultimate classification of given forms of speech ... as dialects or as distinct languages is a very complex matter'. One of the particular problems relating to the recognition of mutual intelligibility involves dialect-chain situations in which mutual intelligibility, of course, is maintained between geographically adjacent communalects,[5] yet over larger distances mutual intelligibility fails. A well-known example of this involves the situation in Germany and the Netherlands, where speakers of local communalects in Amsterdam and Berlin can certainly not understand each other, yet a traveller moving from one communalect to another between these two cities will never encounter mutual unintelligibility.

Although a clear-cut boundary between dialect and language is therefore often not possible, Oceanic linguists have often succumbed to the natural human tendency to operate in terms of discrete entities and clearly defined boundaries rather than allowing for the indeterminacies necessitated by continua. For example, surveys of Vanuatu languages (Tryon 1976) and Solomon Islands languages (Tryon & Hackman 1983) have adopted the traditional lexicostatistical figure of 81% shared cognacy in core vocabulary as representing the boundary between language and dialect, despite clear evidence presented by Wurm and

---

2 But surprisingly difficult to cite.

3 For the most part, my discussion will deal with languages from the Oceanic subgroup of Austronesian, though there will be some reference also to Australian languages and the non-Austronesian languages of Melanesia where this provides relevant supplementary information.

4 The examples discussed by Wurm and Laycock all happen to involve non-Austronesian languages from Papua New Guinea, though the sociocultural contexts are similar enough to what we find for Oceanic languages that their comments can be taken as applying equally to situations of regional diversity within these languages.

5 I will use the term 'communalect' following Wurm and Laycock (1961:132) to refer to a speech form that is indeterminate with respect to separate-language or same-language status.

Laycock (1961) that a variety of phonological and structural factors can also affect mutual intelligibility.

In any case, there is also a real problem in operating with such figures, because different writers comparing essentially the same pairs of communalects can sometimes come up with surprisingly different cognate percentages. In Crowley (1998b:105–106), I point out that a number of lexicostatistical comparisons of pairs of Oceanic communalects by different linguists have produced published cognate figures that vary by as much as 20%. The reasons for this presumably involve, in part, differences of criteria as to what constitute cognates.

Differing degrees of accuracy in raw lexical data have also played a part, especially when comparing figures presented in general surveys against information contained in more detailed studies of particular languages. Davis (1997:25), for example, upwardly revises Tryon and Hackman's (1983) figures for cognate sharing between *Hoava* and *Kusaghe* in Solomon Islands from 77.6% to 89.8% by eliminating errors in their lexical data, which takes this pair of communalects from the status of separate languages to the status of dialects of the same language. One particular source of error involved situations where a particular meaning in both communalects could be expressed by two synonymous forms, yet the lists upon which the lexicostatistical percentages were calculated included only one synonym for *Hoava* and the other synonym for *Kusaghe*, resulting in lower scores for shared cognacy than should have been the case.

Dixon (1997:7) has recently revisited the language/dialect issue by arguing that sociopolitical factors can be factored out of the equation, after which he claims — *contra* Wurm and Laycock — that '... it is generally not a difficult matter to decide whether one is dealing with one language or with more than one in a given situation'.

Empirical verification of the language or dialect status of two speech forms can, Dixon says, be tested by giving people spoken or written passages and then administering comprehension questions (allowing for differences of pronunciation). Comprehension levels above the 80–90% range, he argues, would then be deemed to constitute dialects of a single language.

Unfortunately, Dixon does not attempt to discuss any of the difficulties that would inevitably arise with the administration of such tests. In diglossic situations, where one variety is likely to be considered inappropriate for use in a testing situation, any attempt to apply a test will inevitably produce biased results (Fasold 1984:153), as may turn out to be the case, for example, regarding 'Standard Fijian' and local communalects. Language-testing specialists already have enough difficulty deciding what constitute legitimate testing procedures, yet Dixon proposes to quantify comprehension, which is inherently difficult to quantify (Nettle 1999:63). In any case, one wonders what the precise basis is for Dixon's particular cut-off point in comprehension scores (and how should we interpret his allowance of a range of 10%?). Wurm and Laycock (1961:132–133), and some other writers, suggest much lower rates of information transfer as representing the boundary between dialect and language, though Dixon does not address the variation between his figures and theirs.

Also, how could one ever expect to administer a comprehension test between two languages in societies where there is either active or passive bilingualism between those languages? In the typically multilingual areas where Oceanic languages are spoken, of course, bilingualism of various kinds is the norm rather than the exception (Wurm and Laycock 1961:136). Wurm and Laycock (1961:136) make the obvious point that even the subject of a discourse may influence mutual intelligibility, and speakers of different communalects will almost certainly find it easier to overcome regional differences when they are listening to speech on a subject where they have overlapping fields of experience than

when they do not. Lippi-Green (1994) points out that mutual intelligibility depends on a whole range of additional non-linguistic factors, such as attitudes, beliefs, and even good will.

Dixon (1980:35–36) offers another test for language versus dialect status: the one-or-two-book test. By this test, he argues that if it is more convenient for a linguist to write a single grammar of two communalects — with notes on regional differences — then it is a question of dialects of a single language. If, on the other hand, it is necessary to write two separate grammars, then it is clearly a question of two separate languages. Of course, this test does not tell us how many notes of regional differences we will need to accumulate before a separate grammar is warranted. Obviously, different writers — or publishers — might be prepared to operate according to different aesthetic judgements when making this kind of decision, which means that the division between language and dialect becomes little more than the personal whim of an academic linguist, or even a publisher who knows nothing about linguistics.

For example, the Sye and Ura languages of Erromango in Vanuatu are clearly separate languages according to the mutual intelligibility criterion, as speakers of Sye cannot understand Ura when the language is played to them on tape.[6] However, structurally there are so many direct parallels between the two that I could have simply copied the files from Crowley (1998a), substituted Ura examples for the Sye ones, and made a few amendments to the text to account for the relatively small number of additional differences in producing Crowley (1999a).

Situations like this where the patterns of one language are largely mapped morpheme-by-morpheme onto those of another language, but with partly (and sometimes even completely) different forms, are certainly not unique — see Thurston (1987) for a description of what we find in parts of New Britain for example, and Grace (1981:157–159) for reference to a similar situation in New Caledonia — and one wonders how the one-or-two-book test would be applied here. A shared grammatical text with separate examples is far from impossible in such cases, even though the mutual intelligibility criterion indicates that we are dealing with separate languages.

Dixon (1997:8) acknowledges the existence of dialect chains, for which he concedes that 'fairly arbitrary' decisions may be needed, though he claims that such situations are 'rather rare'. This is a somewhat surprising claim, given that it is widely known that the entire Romance-speaking area of Europe, as well as many parts of Germanic-speaking Europe, constitute gigantic dialect chains (Crystal 1987:25). Also, the speech form that Dixon (1988) described in Fiji belongs to what most would regard as a Fiji-wide dialect chain, or possibly one of two dialect chains (Geraghty 1983:277), despite Dixon's attempt to reduce Fiji to a straightforward two-language situation, with each language having 'a considerable number of dialects' (Dixon 1988:1). Other dialect chains are encountered among Oceanic languages in some parts of Papua New Guinea (Wurm & Laycock 1961:137), including, for example, Central Province (Pawley 1975:10), as well as the Caroline Islands of Micronesia (Lynch 1998:27).

Dixon (1980:37) may be correct in claiming that in Australia — apart from the Western Desert and Central/South Queensland — there were no dialect chains at all. However, he does not consider at least the possibility that in other parts of the continent, earlier dialect chains may have been obliterated soon after European contact by the complete loss (often

---

6 All Ura speakers, however, are bilingual in Ura and Sye, so the mutual intelligibility test could not be applied with them.

even without substantial records) of many intermediate speech varieties, or that the establishment of mission stations and government reserves resulted in considerable amounts of dialect levelling. The island of Erromango, for example, was described by earlier observers as having been linguistically much more like modern-day neighbouring Tanna, with its gradual transitions between varieties. However, major linguistic attrition has brought about the loss of all transitional varieties on Erromango, resulting in now quite clear-cut language boundaries (Lynch 1983:5).

## 3 Traditional naming practices

I now propose to investigate the dialect/language issue from a somewhat different perspective. Rather than adopting the traditional academic criterion of mutual intelligibility, I propose to look at the issue from an indigenous perspective, as reflected in how speakers of Oceanic communalects talk about their own ways of speaking, those of other groups, and also regional diversity within their speech communities. In particular, I will concentrate on the traditional naming of communalects by speakers of Oceanic languages, as well as other conventionalised lexical expressions used in talking about linguistic diversity. However, while I propose to describe the main patterns of naming that are encountered among Oceanic communalects, there will be some reference to non-Oceanic communalects as well.[7]

### 3.1 Talking about diversity

Speakers of Oceanic languages seem generally to be aware of at least some aspects of regional variation within their areas of mutual intelligibility. Geraghty (1983:18), for example, indicates that even very young speakers of Fijian communalects are typically very much aware of even small linguistic differences between their own speech and that of others, and that people generally have a good idea of how far their own communalect extends.

However, the precise characterisation of linguistic differences often involves a concentration on particular kinds of differences while ignoring others. This observation sits well with my own observations of linguistic diversity on Paama and Erromango, which suggest that there is often an element of exaggerating differences by stereotyping. For instance, Erromangans will typically describe a southern dialect of their language as having *h* in words that have *s* in the northern dialect, but the situation is far more complicated than this, as described in more detailed in Crowley (1998c).

People are also able to recognise and talk about varying degrees of difference between mutually unintelligible varieties. The Paamese, for example, will normally say that the people of neighbouring Southeast Ambrym speak a 'different language', which they cannot understand. However, I have on occasion also heard people say that the Southeast Ambrymese speak the 'same language' as they do, though this has always been in the context of comparing Southeast Ambrymese with languages from other parts of Ambrym. Even a linguistically fairly naive observer soon comes to realise that, despite the mutual unintelligibility between Paama and Southeast Ambrym, there is a large number of common

---

7 Additional observations for Australian communalects can be found in Dixon (1980:40–43). Foley (1986:22–29) makes some observations concerning the naming of non-Austronesian communalects in the New Guinea area. For the sake of stylistic convenience, I will describe all of these patterns using the present tense, though it should be kept in mind that some of the traditional patterns described in this section actually refer to communalects that have become extinct, or which are moribund.

individual words that are either the same — or at least very similar in shape — in contrast to the much more divergent vocabularies of the other languages of Ambrym.

The fact that speakers of Oceanic languages clearly talk about linguistic diversity at different levels of generality depending on the context means that different people can end up saying quite different things about exactly the same linguistic situation. John Lynch (pers. comm.), for example, reports that some people on Tanna claim that there is just a single language on the island, while he has heard one person claim that there are as many as twenty-eight languages, and other people have offered various figures between these extremes. Linguists have tended to recognise around three to five languages based primarily on the criterion of mutual intelligibility, and there are no armies or navies on Tanna to provide a more definitive answer (Lynch 1978:719).

### 3.2 Unnamed dialects, named languages

Some communalects in the Pacific have names which function exclusively as language names in the traditional sense described in §2. Many such names are completely unanalysable, such as *Raga*[8] of Pentecost, *Nakanamanga* of Nguna and *Ura* of Erromango, all spoken in Vanuatu. Sometimes, emblematic words within a particular language are chosen as the basis for a language name. For example, on Erromango the various first person singular possessive pronouns ('my') are used also metalinguistically as language names such as *Enyau*, *Aryau*, *Sorug* (Crowley 1997).

Other language names may represent some kind of compound, which may be descriptive in some way of how a people are characteristically seen as speaking. For instance, the name of the *Guugu Yimidhirr* language from north Queensland involves *guugu* 'language' as the initial element, while the second element, *Yimidhirr*, derives from *yimi* 'this' and *-dhirr* 'having'. The name therefore literally means 'language with *yimi* (for 'this')' (Dixon 1980:42). Compound language names sometimes also express some kind of ethnocentric judgement about that group's own way of speaking, for example *Tinata Tuna*, literally 'true language', spoken in the Rabaul area of Papua New Guinea (Lynch 1998:40).

Even very small languages in the Pacific can be expected to exhibit some degree of geographical heterogeneity, and with some languages the diversity can be considerable. *Tinata Tuna*, for example, is spoken over a substantial area of northern New Britain, and it has one of the largest speaking populations in Papua New Guinea today. However, while there is a considerable amount of regional diversity within this language, I am not aware that any of these local dialects are themselves separately named. All are, therefore, equally referred to simply as *Tinata Tuna* 'true language' by their speakers.

Some languages may even have more than one name. For example, *Ura* and *Aryau*, both of which have already been mentioned, are synonymous names for the same language in Vanuatu, one of which is uniquely a language name (*Ura*), while the other is derived from the word in that language for 'my' (*Aryau*). Sometimes one name may be used by speakers of their own language, while other names may be used by speakers of neighbouring languages. For instance, speakers of the *Angkamuthi* language of Cape York in Australia refer to their language by that name, whereas neighbouring groups to the south call their

---

8 Since the precise phonemic shape of language names is not germane to the overall discussion, names are here presented either in the local orthography, or as in the title of the major linguistic description, despite the fact that there are sometimes considerable differences in the phonemic interpretation of particular orthographic symbols.

language *Ngkamuthi* (reflecting the regular loss of initial vowels in their languages), and Torres Strait people refer to it as *Kuta* (Crowley 1983:310–311).

### 3.3 Unnamed dialects, unnamed languages

It is not uncommon in other places for there to be no indigenous names for communalects at all (Lynch 1998:40). In much — though by no means all — of northern and central Vanuatu, as well as in most of Polynesia and Micronesia, there are typically no lexical items that have unique reference as communalect names in traditional usage. In such cases, when people need to refer specifically to their language, they will typically refer to it in terms such as 'the language of such-and-such a place' if that language is associated exclusively with a well-defined geographical location.

The people of the island of Paama in Vanuatu, for instance, having no separate name for their own language, refer to it in contrast to anybody else's language by saying *selūsien tenout Voum*, which literally means 'language of Paama'. Alternatively, they can refer to it unambiguously within their own speech community as *selūsien orer*, which literally means 'our (plural inclusive) language'. When they are speaking in Bislama to somebody from another language area, they can refer to their language unambiguously as *lanwis blong mifala*, literally 'our (plural exclusive) language'.

The Paamese have also incorporated the Bislama word *lanwis* 'language' — in the shape *lanūs* — into their vernacular. This word enters into a highly specialised grammatical construction as a postverbal nominal complement to the intransitive verb *selūs* 'speak' along with other language names. For example:

(1) *Naselūs Veranis.*
'I speak French.'

(2) *Koselūs Inglis.*
'You speak English.'

(3) *Niselūs Pislama.*
'I will speak Bislama.'

(4) *Kiselūs Lanūs*
'You will speak Paamese.'

It should be pointed out, however, that borrowed *lanūs* can, in an appropriate context, refer to any vernacular, though a vernacular other than Paamese is more likely to be accompanied by some additional specification for the location of its speakers, e.g. *lanūs tenout Tanso* 'language of Southeast Ambrym'. Unmodified *lanūs*, on the other hand, is most likely to be interpreted as referring particularly to Paamese.

When someone is speaking Bislama, the word *lanwis* can be used at different levels of generality or specificity, depending on the communicative needs of the context, in a similar way to the relative use of 'same' and 'different' noted in §3.1. For example, a speaker of *Ninde* on Malakula in Vanuatu could speak of his/her vernacular to an outsider simply as *lanwis blong Malakula* 'language of Malakula' if the implied contrast is with any other language from Vanuatu, even though there are nearly thirty other mutually unintelligible languages spoken on the island. Increasing degrees of specific identification could be achieved by referring to *Ninde* as *lanwis blong saot Malakula* 'language of south Malakula',

*lanwis blong Sawes Bei* 'language of Southwest Bay', and finally, to distinguish it from other languages spoken in the multilingual settlement at Southwest Bay, *Ninde*.

Lacking a generic language name certainly does not imply that members of a speech community will necessarily have names for any of the specific regional dialects of that language. Paamese speakers will typically point to a distinction between mutually intelligible northern and southern varieties, though they do not have local names for these. If information about the geographical source of a particular variety is to be expressed, Paamese speakers tend to refer variously to speakers from 'up there' or 'down there', or to refer to the village of origin of a particular speaker.

When describing the differences between northern and southern Paamese, people will often use some fairly impressionistic terminology, referring to people from the north who are said to *vit kotehei* 'speak cut' the language, whereas southerners are said to *lehei* 'pull' the language.[9] In other situations involving regional diversity, different varieties are sometimes impressionistically described instead as 'heavy' or 'light' (Geraghty 1983:18), though these do not always correspond to well-defined linguistic features. On the other hand, while speakers of Erromangan typically stereotype linguistic features that are associated with geographical areas as I have already indicated, they have no conventionalised metalinguistic terminology to refer to regional differences within their language.

### 3.4 Named dialects, unnamed languages

It is also fairly common for what might be considered as geographical dialects of a single language — purely on the grounds of mutual intelligibility — to have no accepted generic language name, but for each of the local dialects to have a name of its own. This kind of situation is encountered in parts of the Solomon Islands, where, for example, separately named *Mbatambana*, *Katazi*, *Sengga*, *Lömaumbi* and *Avasö* are all mutually intelligible varieties spoken on central and eastern Choiseul (Tryon & Hackman 1983:27). Davis (1997:22) indicates that speakers of mutually intelligible but separately named *Hoava* and *Kusaghe* in the Solomon Islands traditionally use these terms alternately as names of the specific dialects, or as names for the language as a whole. However, she indicates that when people need to unambiguously refer to their language as a whole in contrast to other languages, they coordinate both names into a single phrase.

Linguistically diverse Fiji also falls into this category (Geraghty 1983). Although we do not have enough data to be certain, it is also possible that separately named *Enyau* and *Sorug* on Erromango in Vanuatu could also be considered linguistically as having been separate but mutually intelligible varieties for which there was no overall name (Crowley 1997:47–48). Lynch (1978:719) also indicates that the range of recognisable communalects on Tanna are grouped into eight named varieties.

Northern Malaita is another area in which local dialects have names, but there are no generic language names. Siegel (1987:219–220) indicates that people of this area regard language as an important element in group identification, and the dialect names often seize upon local linguistic habits in the derivation of names. For instance, *Baelelea* is a compound involving *bae* 'say' and *lelea*, which reflects the local predilection for reduplicating the word *lea* 'go' as *lelea*.

---

9 This terminology reflects the fact that by and large, the southern dialect is phonologically more conservative, while words in the northern dialect have generally undergone various processes of phonological deletion or assimilation.

### 3.5 Language names and ethnonyms

In some societies, the name of an ethnic group and the language of that group may be quite different. Anglo-Celtic New Zealanders, for example, can be referred to by the ethnonym *Pākehā* while their language is English, never *Pākehā-ese*. Such situations, however, are not all that common, and it is far more common for there to be some kind of formal relationship between language name and ethnonym, with either the language name being derived from the ethnonym, or the ethnonym being derived from the language name.

Dixon (1980:40–41) points out that ethnic groups in indigenous Australia (often referred to as 'tribes') frequently derive their ethnonyms from their language name by means of a derivational suffix of some kind. Thus, the name of the *Yidinyji* people of northern Queensland derives from the language-name *Yidiny* by means of the suffix *-ji*, which expresses the meaning of 'having'. The *Yidinyji* are therefore those people who have (i.e. speak) *Yidiny*. Yet other tribes have names that appear to be formally derived from their language name, though by means of otherwise unrecognisable morphemes, such as with the *Jirrbalngan* people, who speak *Jirrbal* (Dixon 1980:42).

There are other cases, however, where it appears that it is the ethnonym which is basic, and the language is referred to in terms of being the speech of a particular ethnic group (similar to how European language names typically derive from the names of nations or regions, e.g. *Russia* > *Russian*, *Italy* > *Italian*, *Galicia* > *Galician*). Thus, the word *Motu* primarily refers to the people who live along the coast around Port Moresby, while their language has no distinct name of its own (Lynch 1998:40). It is necessary to refer to the language by means of the descriptive phrase involving *gado* 'language'.

It should be pointed out that boundaries of ethnicity and language clearly do not need to coincide. Dixon (1980:35) points out that originally the 600 or so separate 'tribes' of Australia spoke only about 200 separate languages. Foley (1986:23) also points out that merely sharing a common language does not guarantee that people will necessarily identify as a political entity, as there are a number of villages in the Karawari area of the Sepik in Papua New Guinea where people speak a series of mutually intelligible varieties, yet there is little sense of either a single linguistic or political community. In fact, people's non-linguistic links seem to be closer in many cases to neighbouring villages where people speak mutually unintelligible varieties (though obviously with extensive patterns of bilingualism operating throughout the area). Nineteenth-century sources also indicate that the main traditional political groupings on Erromango did not correspond closely to linguistic boundaries, leading to a situation where speakers of the same language could belong to separate political groupings, while speakers of different languages could belong to the same political entity (Spriggs & Wickler 1989).

## 4 Postcontact language names

As far as possible, the discussion in §3 has been presented to describe the kinds of naming patterns that prevailed among speech communities prior to colonial contact. Since that time, of course, much has changed, including in some cases the ways that people refer to their languages. Massive depopulation and major movements of people have resulted in considerable loss of original linguistic diversity in some areas. This has happened most notably in Australia, where either most of the original languages have completely disappeared, or their loss seems imminent. However, it is possible that there has also been some unrecorded loss of linguistic diversity among Oceanic languages due to depopulation in

some parts of Vanuatu such as parts of Malakula, and several languages have certainly been lost on Erromango due to depopulation (Crowley 1997).

In such circumstances, what was once a more diverse area may have experienced changes in the practices associated with language naming. Crowley (1997), for example, documents the replacement of original *Enyau*, *Sorug*, *Utaha* and *Ura* on Erromango with a single viable modern language, which appears to be basically a koine based on original *Enyau* and *Sorug*. However, neither of these names is now used, and no unique language name has been adopted to replace these forms.[10] It seems that the loss of linguistic diversity on the island has eliminated the need for people to make any kind of explicit contrast between the language that they speak today and any other language that was formerly spoken on the island, so there is a reduced need for it to have a distinct name. Thus, the language of Erromango today has changed from being a named language of the type described in §3.1 to an unnamed language, as described in §3.3.

In yet other cases, a previously unnamed language has acquired a name, which has arisen from within the community of its own speakers. The indigenous people of the North and South Islands of New Zealand originally had no separate name for their ethnic group as a whole, or for their language, though they had many local names for the various *iwi* ('tribal') groupings, e.g. *Ngāti Porou*, *Tūwharetoa*, *Ngāpuhi*, *Kai Tahu* etc. (though not their local dialects). With the arrival of substantial numbers of European settlers (who came to be known from very early on by the etymologically obscure word *Pākehā*), the common sense of indigenous, i.e. non-*Pākehā*, ethnicity came to be expressed by semantically extending the original word *māori*, which meant 'plain, ordinary' to become the ethnonym *Māori*, from which is derived the name of the language *te reo Māori* (< *te reo* 'language').

In a similar way, *na vosa vaka-Viti* 'Fijian language' (*na vosa* 'language', *vaka-* 'adjectival derivative', *Viti* 'Fiji') as a language name has been superimposed over the various local named speech varieties. In this case, however, there was also an associated development of a semiartificial written standard based on one of the local varieties that was promoted by nineteenth-century European missionaries.

Original language names have also sometimes been replaced by new names that derive from placenames associated with colonial government or mission resettlement. Lynch (1998:41) reports that the indigenous names for what are now commonly referred to on Tanna in Vanuatu as the *Lenakel* and *Waitsan* languages are seldom used by local people today. Both of these new language names derive from the names of the settlements *Lenakel* and *Whitesands* that developed as important centres in the colonial era within these two areas.

In yet other cases, an introduced word has filled a gap corresponding to the lack of a traditional language name. Once the original language names were lost on Erromango — as described above — the sole surviving viable language has come to be referred to today in contrast to other languages in Vanuatu as *nam Eromaga*, literally 'Erromangan language', even though *Eromaga* is itself an introduced word (about the source of which there is some dispute).

However, other postcontact language names referred to by Lynch (1998:40–41), Lynch (1994:viii) and Dixon (1980:40–43) have been derived by means of English derivational

---

10 Published descriptions of the language such as Crowley (1998a) refer to it as *Sye*. However, this is simply one of the alternative names for linguistic varieties recorded in nineteenth-century sources which is remembered by a handful of people on the island as a language name. Most speakers of the language no longer use, or even recognise, *Sye* as a language name, hence my inclusion of the form within parentheses in *An Erromangan (Sye) grammar*.

morphology on the basis of the names of places in English where a language is spoken (e.g. *Paama* > *Paamese*, *Erromango* > *Erromangan*, *Fiji* > *Fijian*),[11] or by compounding existing dialect or placenames into a single word (e.g. *Ifira-Mele*, *Aniwa-Futuna*). Yet other newly created language names derive from the choice of a single local dialect name — sometimes arbitrarily — as a convenient 'cover term' for all of the regional dialects involved, such as Davis' (1997) choice of *Hoava* for both mutually intelligible *Hoava* and *Kusaghe*.

It should be pointed out that in most instances the deliberate creation of such names has simply been to allow outsiders to talk about these languages amongst themselves in English. This has involved either academic linguists who have produced grammars and dictionaries, or missionaries who have formulated and implemented educational and evangelistic programmes in these languages. I am aware of very few cases where this kind of externally imposed linguistic labelling has actually impinged on local usage in any way.

Despite the fact that a grammar and a dictionary has been published for 'Paamese', for example (Crowley 1982, 1992), or 'Erromangan' (Crowley 1998a), the way in which the local people name these two languages has not changed in any way. In fact, Lynch (1994) represents a comprehensive compilation of alternative language names for Vanuatu languages, and this is valuable precisely because so many of these competing externally imposed labels have *not* entered general usage. We therefore find, for example, a reference to the locally used language name *Nakanamanga*, along with the following competing externally imposed labels:

> Efate(se), Havannah Harbour, Nguna, Ngunese, North Efate, Sesake, Tongoa(n), and a number of other aliases. (Lynch 1994:36)

These names derive from a range of sources: placenames in English (*Havannah Harbour*, *North Efate*), anglicised placenames derived from indigenous sources (*Efate*), indigenous placenames (*Nguna*, *Sesake*, *Tongoa*), or derivations from indigenous placenames using English morphology (*Ngunese*, *Tongoan*, *Efatese*).

As I pointed out in Crowley (1999b), local people are largely unaware of the existence of published volumes dealing with their languages. I have, for example, deliberately chosen not to distribute copies of Crowley (1982, 1998a) widely to members of the Paamese and Erromangan communities because the way in which their languages are described grammatically renders those books embarrassingly difficult — and sometimes even impossible — to understand. Small numbers of my dictionary of Paamese (Crowley 1992) have been distributed to people on Paama, but this is primarily a Paamese–English dictionary, rather than primarily an English–Paamese dictionary, or a monolingual Paamese volume. For this reason, while the information contained within it may be of some interest, it is likely to be of little practical use to members of the local community. The copies are therefore seldom seen or used on Paama, and the dictionary is likely to be of much greater interest to comparative linguists or other Oceanic specialists.

Another modern language-naming situation that is worth mentioning involves people of Aboriginal ancestry in Tasmania, who have a keen interest in seeing community members make greater use of words from their ancestral indigenous languages, though the last speakers of these languages died in the 19th century. Tasmania was once linguistically fairly diverse, with possibly as many as a dozen distinct languages spoken there (Crowley & Dixon 1981). However, the records of each of these languages on their own is so poor that the attempt at linguistic revival has been based on the idea that words from the various recorded

---

[11] The indigenous Paamese word for the island is *Voum*, while the Erromangans call their island *Unelocompne*, and the Fijian word for Fiji is *Viti*.

vocabularies should be pooled together to constitute one collective lexical resource. The language that is being promoted in this way by the Tasmanian Aboriginal Centre is referred to by the locally coined phrase *Palawa kani*, which is based on the recorded words *Palawa* 'Aboriginal person' and *kani* 'talk'. These words have been compounded according to the English pattern, in the absence of any substantive grammatical information which would indicate how — or indeed even if — compounds could be formed in the original languages.[12]

With regard to this latter case, it is interesting to note that outsider academics have again attempted to create a new language name, apparently without realising that members of the local community had independently coined their own preferred name. Thus, Mühlhäusler (1996:128) names and identifies their language as *Neo-Tasmanian*, possibly on the analogy of Hall's (1959) unsuccessful term *Neo-Melanesian* for Melanesian Pidgin.

It is worth noting, however, that in similar situations of complete loss of the original language such as we commonly find in southeastern Australia, present-day exclusively English-speaking Aboriginal people are expressing increased interest in their traditional past. With regard to language, this knowledge is often derived exclusively from older, and often phonetically unreliable, written sources. The result is that original language names such as Nganyaywana, Djangati and Gumbaynggir are regarded with suspicion, with people referring instead to the languages instead as Aniwan (pronounced as 'Anna won'), Dungguti and Koombanggee. These names are either spelling pronunciations based on older written sources, or anglicisations of the original language name. In cases such as 'Aniwan', the older poorly recorded language name appears to have given a level of reverence which derives solely from its long tradition of appearing in writing, despite its lack of empirical validity.

## 5 New languages

Since the advent of the colonial era in the Pacific, we have seen the emergence of a wide range of new languages that have arisen in situations of multilingual contact, i.e. new pidgin/creole varieties (Lynch 1998:220–236). I propose to examine only the major themes involved in the naming of such languages, not because they relate to the issue of the difference between dialect and language in Oceanic languages, but because it relates to a broader issue arising out of dialect and language naming that I propose to address in §6.

In some cases, the speakers of these new languages have provided their own names, which they have derived from a variety of sources. The name of French-lexifier *Tayo* in New Caledonia reportedly derives from a word in Tahitian meaning 'friend' (Hollyman 1983:133–136) and younger speakers of the language have recently adopted this as an in-group term which identifies their language (Ehrhart 1993:51). The English-lexifier creole spoken in the area of Cape York and Torres Strait is referred to locally as *Broken*, which clearly derives from 'broken (English)'.[13]

These pidgins and creoles have sometimes also acquired vernacular names which people use when speaking about the languages within their own communities. For instance, *Bislama* is referred to by Erromangans as *Nam Ilvucteven*, which literally means 'between language', reflecting its role as a language of contact between Melanesians and English- or French-

12 An interesting Austronesian — though not Oceanic — parallel involves some Netherlands-born Moluccans who are endeavouring to teach themselves 'Moluccan' from published sources which derive from a variety of linguistic sources, resulting in an artificial amalgam of several different Moluccan languages (John Bowden, pers. comm.).

13 This name presumably reflects the community's earlier negative stereotyping of their own language.

speaking Europeans. People from southern Tanna who speak *Kwamera* but no Bislama tend to refer to it by the phrase *nagiarian sei pitoga*, literally 'language of the foreigners'.

In other cases, the group who initiated a particular name has been obscured by time. For instance, while *Bislama* derives its name ultimately from Portugese *bicho de mar* — referring to the sea slugs that represented a major trading commodity in the formative years of the language around the middle of the 19th century (Crowley 1990:26–33) — there is no way of knowing whether that word became a language name at the behest of its earliest Melanesian speakers or of European traders and labour recruiters at the time. However, the active involvement of local people is suggested by the somewhat archaic alternant *Bislaman*, which presumably originated analogically on the basis of *Inglisman* 'English person' and *Franisman* 'French person'.

In Papua, the colonial police force in the late 19th century was centrally involved in the formation and spread of a pidginised form of Motu, which is the vernacular language spoken around Port Moresby. People generally refer to both pidginised and vernacular Motu as *Motu*, though if a distinction is to be encoded, pidgin Motu is typically referred to as *Motu*, while vernacular Motu is referred to as *Motu korikori* 'true Motu'.

As a result of the circumstances in which this pidgin language originated, it came to be referred to in English — by both Europeans and educated Papua New Guineans — as *Police Motu*. In the lead-up to the independence of Papua New Guinea in 1975, there was something of a Papuan separatist movement which was in part linguistically inspired, and some *Police Motu*-speaking Papuans were concerned about being dominated by *Tok Pisin*-speaking New Guineans[14] from the early 1970s.

Objecting to *Tok Pisin* as a language of obvious colonial origins, it was claimed — incorrectly as it turned out (Dutton 1985) — that the Papuan lingua franca originated in the precolonial era at the time when the well-known *hiri* trading expeditions involved local people in major exploits of organisation and long-distance ocean voyaging. This prompted a move among intellectuals to rename *Police Motu* as *Hiri Motu*, and this new — though historically inaccurate — name has become well-established, at least among better educated speakers of the language (Dutton 1985:127–128).

Europeans — whether academic linguists, missionaries or government officials — have from time to time also come up with their own names for some of the pidgins and creoles of the region, for a variety of reasons. For instance, in order to refer generically to *Tok Pisin*, *Pijin* and *Bislama* as a single language with three mutually intelligible national varieties spoken in Papua New Guinea, the Solomon Islands and Vanuatu respectively, the term *Melanesian Pidgin* has been widely adopted in English (e.g. Holm 1989:526), while others have adopted alternative idiosyncratic creations such as *Neo-Melanesian* (Hall 1959) or *Bislamic languages* (Sankoff 1996:421). For the most part, however, these are terms which have enabled outsiders to discuss various aspects of these languages in English, and they have not had any impact whatsoever on local usage, or even general academic usage.

The only situation of which I am aware where a name introduced by an academic has actually been incorporated into local usage involves the *Kriol* language of the Roper River area of northern parts of Australia. This language was originally variously named by its speakers as *Pijin*, *Pijin Ingglij* or *Blekbala Ingglij*. However, with moves since the 1970s towards the development of vernacular literacy and vernacular education in the Northern Territory,[15] a writing system and accompanying reading materials were developed in this

---

14 *Tok Pisin* is the English-lexifier contact language that is widely used as a lingua franca in the northern part of Papua New Guinea.

15 Such programmes, however, are currently facing the possibility of being axed.

language, and those who used the language as a medium of instruction in schools were deliberately introduced to the term *Kriol*, and it has apparently spread from there (Sandefur 1979:7–8).

## 6 Language-naming as linguistic trespass

The discussion in §3 shows that there is a very wide variety of traditional practices to be encountered among Oceanic languages with regard to the naming of communalects, ranging from the naming of only local geographic varieties, to the naming of 'languages' corresponding closely to the traditional linguistic definition of the term, and also the lack of any explicit lexicalised communalect names (though usually with some conventionalised means of referring to languages, but not local dialects).

Indigenous language-naming practices have clearly not been static, with older language names sometimes disappearing since colonial contact, and in other cases new names being introduced by local people into their own languages (§4). Yet other language names have been deliberately created by academics or missionaries, though in most cases this has been simply to allow them to talk (or write) in English about the linguistic situations in which they were operating (§4, §5).

However, Mühlhäusler (1996) claims that prior to colonial contact the notion of 'language' in the sense described in §2 was absent from the Pacific, and that:

> ... the concept of 'a language' is brought into existence by this process [of colonialism]
> ... Indigenous conceptions of language thus have given way to European concepts ...
> (Mühlhäusler 1996:53–54)

But while Mühlhäusler asserts that traditional Pacific societies had no concept of 'language', he does not carry out any detailed study of the Pacific similar to the survey in the present paper by way of verification.

According to Mühlhäusler, languages are so non-existent in what he refers to as the 'linguistic ecologies' (Mühlhäusler 1996:238) of the Pacific that even the word 'language' has no translation equivalent in the languages of the region. However, there are in fact *many* Oceanic languages which have words which clearly express precisely this meaning. The following represent a random selection, and Siegel (1997:228–229) points out that there are other languages which can be added to this list:

| (5) | | |
|---|---|---|
| | Erromangan | *nam* |
| | Paamese | *selūsien* |
| | Southeast Ambrym | *seppinien* |
| | Fijian | *vosa* |
| | Māori | *reo* |
| | Tinata Tuna | *tinata* |

In fact, of all of the Oceanic languages that I have come into contact with, either through my own fieldwork or library research, I cannot think of *any* language that does *not* have a word for 'language'.[16]

Of course, many of these words express other meanings as well, such as 'talk' or 'utterance', and some are nominalisations of the verb meaning 'talk'. However, the fact that these words do not uniquely translate the English word 'language' should obviously not be

16 In many cases, this is the *only* metalinguistic term that I have encountered in Oceanic languages.

taken as an indication that speakers of these languages are not familiar with this concept. According to my *Cassell's German and English dictionary*, the word *Sprache* translates not only as 'language', but also 'speech', 'diction', 'parlance', 'voice', 'accent', 'style' and 'discussion'. However, this clearly does not mean that speakers of German do not have a word corresponding to 'language' in English.

Despite this kind of evidence, Mühlhäusler finds the notion of countable and nameable languages to be so objectionable that in his response to in Siegel's (1997) criticisms on this point, he attempts at one point to disown the term by replacing 'language' with the phrase 'Siegel's metalanguage' (Mühlhäusler 1998a:219). It is interesting to note, however, that Mühlhäusler (1998b) himself refers elsewhere to individual languages using the term 'language' a total of eighteen times, in contexts such as the following:[17]

> ... three languages with which I have had recent experience, Norfolk, Milne Bay English, and South Australian Nunga English (Mühlhäusler 1998b:357)
>
> ... a language which occupied four years of my undergraduate studies, Afrikaans (Mühlhäusler 1998b:357)

In fact, not only does he himself use this metalanguage here, but he also 'names' (and 'counts') Norfolk, Milne Bay English, South Australian Nunga English and Afrikaans.

Mühlhäusler goes considerably further than just saying that 'languages' did not exist in the Pacific, and that the languages of today are essentially colonial creations. He claims, in fact, that the academic tradition of naming languages itself constitutes a dangerous colonial practice:

> ... [T]he identification of languages and their subsequent naming is far from being an act of objective description, and it can constitute a very serious trespass on the linguistic ecology of an area. The very view that languages can be counted and named may be part of the disease that has affected the linguistic ecology of the Pacific ... (Mühlhäusler 1996:5)

It can be seen from this quotation that not only is Mühlhäusler denying the existence of 'languages', but that he sees any attempts to identify, name and count them as having deliterious effects on the languages of the Pacific, with linguistic diversity already having been lost, and a considerable amount of additional loss of diversity soon to explode upon the scene as a result.

Not only is Mühlhäusler in serious contradiction to his own stated position on the issue of naming and identifying languages by naming languages himself, but his fundamental assumptions that speakers of Pacific languages do not operate with the notion of 'language', and that people in the greater Pacific do not 'name' and 'identify' their own languages, demonstrate a serious level of unfamiliarity with the facts for at least many parts of the Pacific. In fact, many Pacific languages have names which function purely as language names (as noted in §3.1), a situation which contrasts with most European languages, where names are typically related in some way to what is primarily either an ethnonym or the name of a nation-state.

The vast majority of acts of language-name bestowal that I have described in this paper have originated from within indigenous speech communities, so they can hardly constitute the externally imposed linguistic trespasses about which Mühlhäusler aims to warn us. The externally imposed language names to which I refer in §4 and §5 have for the most part not influenced local usage in any significant way, as these names have almost exclusively been

---

17 I have not counted quotes such as 'the linguistic nature of Norfolk' (Mühlhäusler 1998b:358), though given the semantic relationship to 'language', this should probably have been included in my count. I also ignored the use of the term 'language' when Mühlhäusler was citing or paraphrasing the words of others.

used to enable academics and other outside professionals to talk about the language situation in the region amongst themselves in English in an almost exclusively academic context.

My basic point, then, is that both Dixon and Mühlhäusler are guilty of drastically oversimplifying a complex issue. Dixon oversimplifies by attempting to argue that the terms 'language' and 'dialect' can be defined asocially and apolitically, whereas Mühlhäusler oversimplifies by claiming that the distinction between the terms is nothing more than a colonial invention. I think that it would reflect greater wisdom to stick with the words of Wurm and Laycock of forty years ago: 'the ultimate classification of given forms of speech ... as dialects of a distinct language is a very complex matter'.

## References

Crowley, Terry, 1982, *The Paamese language of Vanuatu.* Canberra: Pacific Linguistics.

—— 1983, Uradhi. In R.M.W. Dixon and Barry J. Blake, eds *Handbook of Australian languages*, vol. 3, 306–428. Canberra: The Australian National University Press.

—— 1990, *Beach-la-mar to Bislama: the emergence of a national language in Vanuatu.* Oxford: Clarendon Press.

—— 1992, *A dictionary of Paamese.* Canberra: Pacific Linguistics.

—— 1997, What happened to Erromango's languages? *Journal of the Polynesian Society* 106(4):33–63.

—— 1998a, *An Erromangan (Sye) grammar.* Oceanic Linguistics Special Publication No.27. Honolulu: University of Hawai'i Press.

—— 1998b, A salvage sketch of Nati (Southwest Malakula, Vanuatu). In Darrell Tryon, ed. *Papers in Austronesian linguistics* No.5, 101–148. Canberra: Pacific Linguistics.

—— 1998c, The voiceless fricatives [s] and [h] in Erromangan: one phoneme, two, or one and a bit? *Australian Journal of Linguistics* 18(2):149–168.

—— 1999a, *Ura: a disappearing language of Southern Vanuatu.* Canberra: Pacific Linguistics.

—— 1999b, The socially responsible lexicographer in the Pacific. *Journal of Multilingual and Multicultural Development* 20(1):1–12.

Crowley, Terry and R.M.W. Dixon, 1981, Tasmanian. In R.M.W. Dixon and Barry J. Blake, eds *Handbook of Australian languages,* vol. 2, 394–421. Canberra: The Australian National University Press.

Crystal, David, 1987, *The Cambridge encyclopedia of language.* New York: Cambridge University Press.

Davis, Karen, 1997, A grammar of the Hoava language, Western Solomons. Unpublished PhD thesis, University of Auckland.

Dixon, R.M.W., 1980, *The languages of Australia.* Cambridge Language Surveys. Cambridge: Cambridge University Press.

—— 1988, *A grammar of Boumaa Fijian.* Chicago: The University of Chicago Press.

—— 1997, *The rise and fall of languages.* Cambridge: Cambridge University Press.

Dutton, Tom, 1985, *Police Motu iena sivarai.* Port Moresby: University of Papua New Guinea Press.

Ehrhart, S. 1993, *Le créole français de St-Louis (le tayo) en Nouvelle-Calédonie.* Hamburge: Buske.

Fasold, Ralph, 1984, *The sociolinguistics of society*. Oxford: Basil Blackwell.

Foley, William A., 1986, *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.

Geraghty, Paul, 1983, *The history of the Fijian languages*. Oceanic Linguistics Special Publication No.19. Honolulu: University of Hawai'i Press.

Grace, George, 1981, *An essay on language*. Columbia (South Carolina): Hornbeam Press.

Hall, Robert A. Jr., 1959, Colonial policy and Neo-Melanesian. *Anthropological Linguistics* 1(3):22–27.

Hollyman, K.J., 1983, Datations et documents. *Observatoire du français dans le Pacifique: Etudes et documents* 1:13–144.

Holm, John, 1989, *Pidgins and creoles*, vol. 2: *Reference survey*. Cambridge: Cambridge University Press.

Lippi-Green, R. 1994, Accent, standard language ideology, and discriminatory pretext in the courts. *Language in Society* 23:163–198.

Lynch, John, 1978, Proto-South Hebridean and Proto-Oceanic. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings, fascicle 2 (Eastern Austronesian)*, 717–779. Canberra: Pacific Linguistics.

—— 1983, The languages of Erromango. In John Lynch, ed. *Studies in the languages of Erromango*, 1–10. Canberra: Pacific Linguistics.

—— 1994, *An annotated bibliography of Vanuatu languages*. Suva: Pacific Information Centre (University of the South Pacific Library).

—— 1998, *Pacific languages: an introduction*. Honolulu: University of Hawai'i Press.

Mühlhäusler, Peter, 1996, *Linguistic ecology: language change and linguistic imperialism in the Pacific region*. The Politics of Language. London: Routledge.

—— 1998a, A rejoinder to Siegel's review of *Linguistic ecology*. *Australian Journal of Linguistics* 18(2):219–225.

—— 1998b, How creoloid can you get? *Journal of Pidgin and Creole Linguistics* 13(2):355–371.

Nettle, Daniel, 1999, *Linguistic diversity*. Oxford: Oxford University Press.

Pawley, Andrew, 1975, The relationships of the Austronesian languages of Central Papua: a preliminary study. In T.E. Dutton, ed. *Studies in the languages of south-east Papua*, 3–105. Canberra: Pacific Linguistics.

Sandefur, John R. 1979, *An Australian creole in the Northern Territory: a description of Ngukurr-Bamyili dialects* (Part 1). Work Papers of SIL–AAB. Series B, vol. 3. Darwin: Summer Institute of Linguistics.

Sankoff, Gillian, 1996, The Oceanic substrate in Melanesian Pidgin/Creole revisited: a tribute to Roger Keesing. In John Lynch and Fa'afo Pat, eds *Oceanic studies: proceedings of the First International Conference on Oceanic Linguistics*, 421–450. Canberra: Pacific Linguistics.

Schütz, Albert J., 1972, *The languages of Fiji*. Oxford: Clarendon Press.

Siegel, Jeff, 1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji*. Studies in the Social and Cultural Foundations of Language 5. Cambridge: Cambridge University Press.

—— 1997, Review of *Linguistic ecology: language change and linguistic imperialism in the Pacific*. *Australian Journal of Linguistics* 17(2):219–244.

Spriggs, Matthew and Stephen Wickler, 1989, Archaeological research on Erromango: recent data on southern Melanesian prehistory. *Indo-Pacific Prehistory Association Bulletin* 9:68–91.

Thurston, William R., 1987, *Processes of change in the languages of north-western New Britain.* Canberra: Pacific Linguistics.

Tryon, D.T., 1976, *New Hebrides languages: an internal classification.* Canberra: Pacific Linguistics.

Tryon, D.T. and B.D. Hackman, 1983, *Solomon Islands languages: an internal classification.* Canberra: Pacific Linguistics.

Wurm, S. and D. Laycock, 1961, The question of language and dialect in New Guinea. *Oceania* 32(2):128–143.

Crowley, T. "The question of dialect and language in Oceania". In Bowden, J. and Himmelmann, N. editors, *Papers in Austronesian subgrouping and dialectology.*
**PL-563:3-20**. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.3*

# 3 *The pretenders to the Muna-Buton group*

MARK DONOHUE

## 1 Introduction[1]

This paper deals with a set of languages that have been assumed to be members of the so-called Muna–Buton group. I propose a separate subgrouping for some of these languages, linking them to Laiyolo and Kalao spoken on the islands in the far south of South Sulawesi, and Wotu in that province's far north-east; in a sense, this paper presents a case for a position that was noted by Sirk (1988) as a probable solution for the subgrouping puzzle surrounding Wotu, but goes further in defining the ways in which the old Muna–Buton group fails to stand up to scrutiny. Although this paper addresses the question of what does not belong to the Muna–Buton group, it does not attempt to list exhaustively the extent of the group, nor to address the question of subgrouping of the Tukang Besi languages, spoken on the nearby Tukangbesi islands, or Kulisusu, spoken in northern Buton, both of which have been asserted to belong to a subgroup with the other languages of Muna and Buton.

## 2 The Muna–Buton area

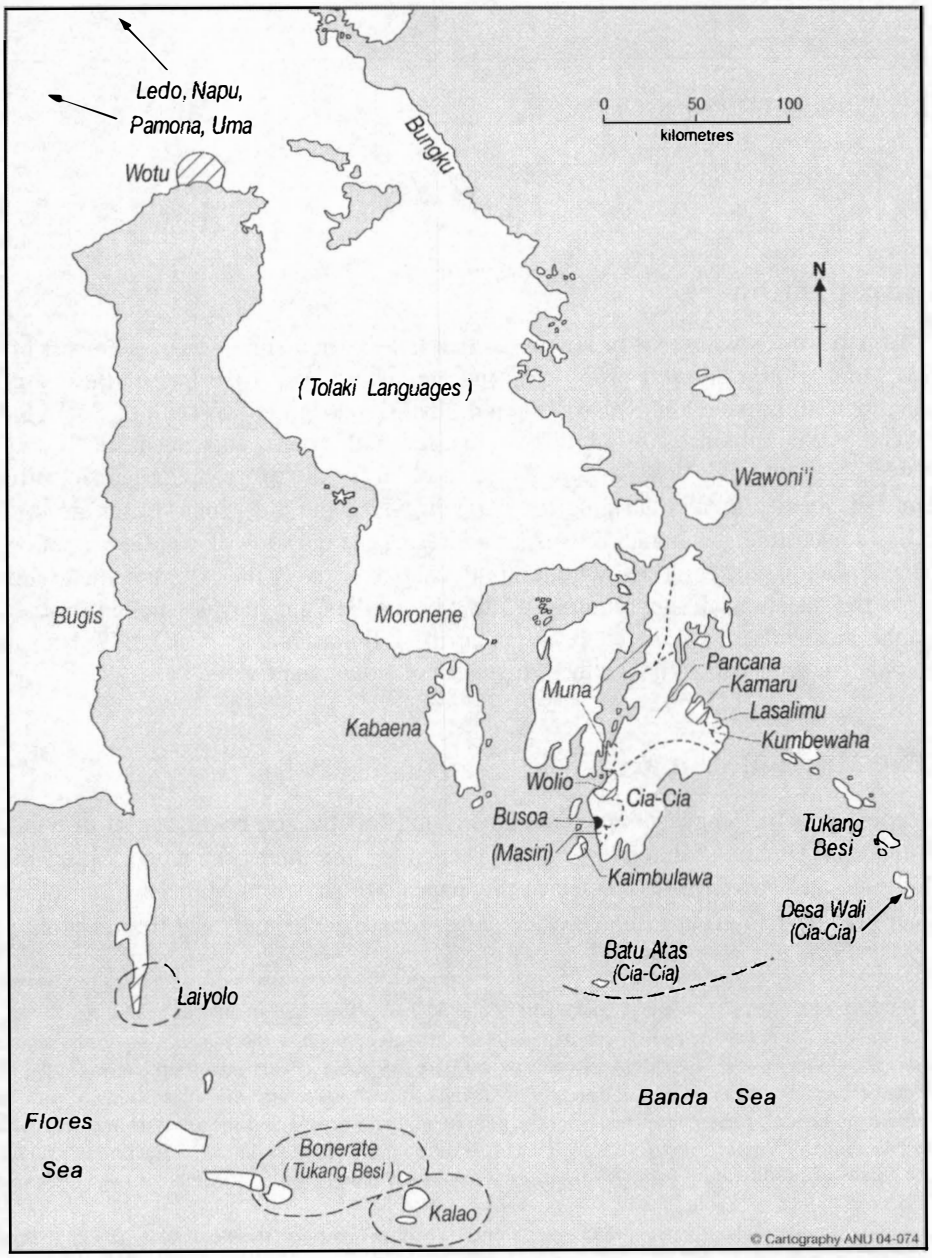
Traditionally the languages spoken on the islands off the southeastern part of Sulawesi on the islands of Muna and Buton have been grouped together under the name 'Muna-Butung'.[2] The region, and languages discussed in this paper, are shown in Map 1. Esser (1938) and Salzner (1960) both grouped together the languages of the islands of Muna, Buton, and the

1 The help of Cathryn Donohue, Chuck Grimes and Malcolm Ross has greatly improved the content of this paper, through either their proof-reading skills or through comments and suggestions on the methodology and how to improve it. In addition, René van den Berg, who shares with me a deep interest in the linguistic prehistory of Southeast Sulawesi, deserves my thanks for encouragement and suggestions before this paper was even begun. I have benefited from the help of all these people, but the analysis and any faults in it, either direct or implied, rest with myself. This is particularly apparent in the bibliography, which reflects the state of the literature in 1996, when this manuscript was accepted for publication. It has not been updated.

2 'Butung' is the Bugis pronunciation of the name. The local pronunciation is usually [ɓu'tɔn], stressed on the second syllable.

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 21–35.
Canberra: Pacific Linguistics, 2004.

Tukang Besi islands (in Southeast Sulawesi) with the languages in the far south of Sulawesi, namely the southern half of Selayar island and the smaller islands in the Sea of Flores (Kalaotoa, Bonerate, Kalao, Tanahjampea, Kayuadi). Whilst showing some differences in their subgrouping, the borders of Esser's and Salzner's groups are comparable. Esser's classification is given in Figure 1 and Salzner's in Figure 2.




**Map 1:** Languages and language groups referred to in the text

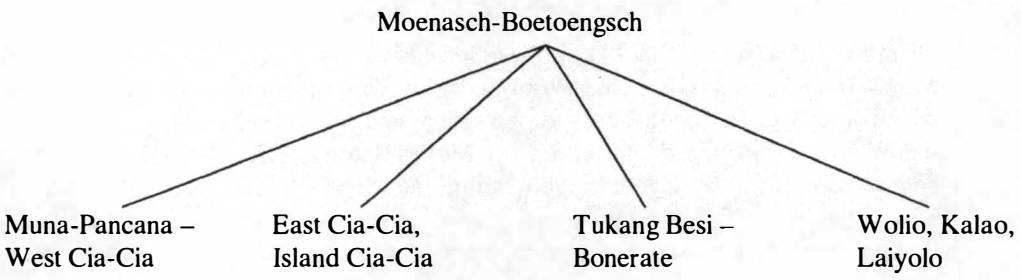



**Figure 1:** Esser's (1938) subgrouping

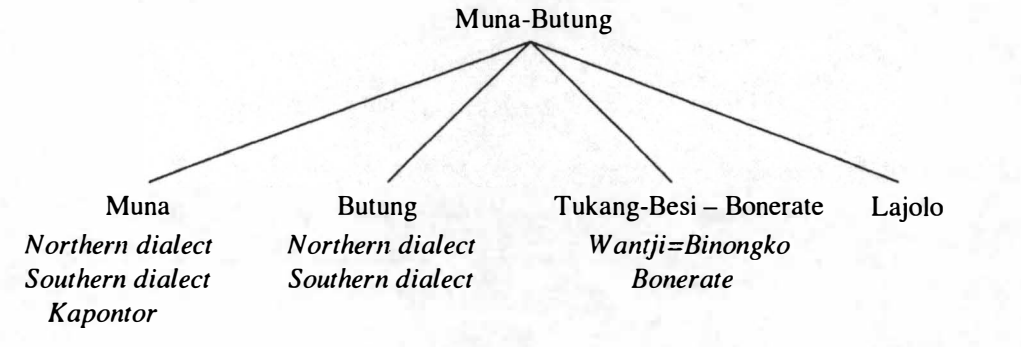



**Figure 2:** Salzner's (1960) subgrouping

The other subgrouping hypotheses that have been proposed are outlined below; in all cases, the subgrouping is by assertion or declaration, with no evidence or methodology presented to enable us to evaluate the proposal.

- Anceaux (1978:281) stated his conviction that '... Adriani's Muna–Buton group has to be reformulated', and recognised subgroupings consisting of Wolio, Lasalimu and Kamaru on the one hand and Cia-Cia, Pancana and Muna on the other. Writing about Tukang Besi, he noted that it '... scores relatively low with all the others'. Nevertheless, he writes in the same paragraph that '... there is reason to believe that all the languages of this area [including Tolaki and Bungku — MD] belong to one subgroup ...'. His subgrouping is given in Figure 3.
- Bhurhanuddin (1979) did not challenge this grouping, and appears to have implicitly divided the languages into five groups, separating Wakatobi (=Tukang Besi), Wolio, Kamaru and Lasalimu from his Muna–Pancana–Cia-Cia group. Bhurhanuddin speculated (as did Salzner (1960)) that Tukang Besi forms a subgroup with Bonerate in the Sea of Flores, but had no data from the language from which to draw conclusions.

- Kaseng et al. (1987) grouped Wolio and Kamaru together, but failed to subgroup any of the other languages that they surveyed in the Muna–Buton area.
- Grimes and Grimes (1987), although only incidentally concerned with the Muna–Buton group, compared Wotu and Laiyolo with Buton [=Wolio (C. Grimes, pers. comm.)] of Salzner's Muna–Buton group, and found evidence to '... substantiate tentatively classifying Wotu within the Muna–Buton Stock' (1987:63), and to '... follow Salzner in classifying Laiyolo within the Muna–Buton Stock' (1987:60). The idea of the Muna–Buton group now extended northwards to include Wotu, a language that has been the source of much disagreement as to its genetic classification amongst scholars.

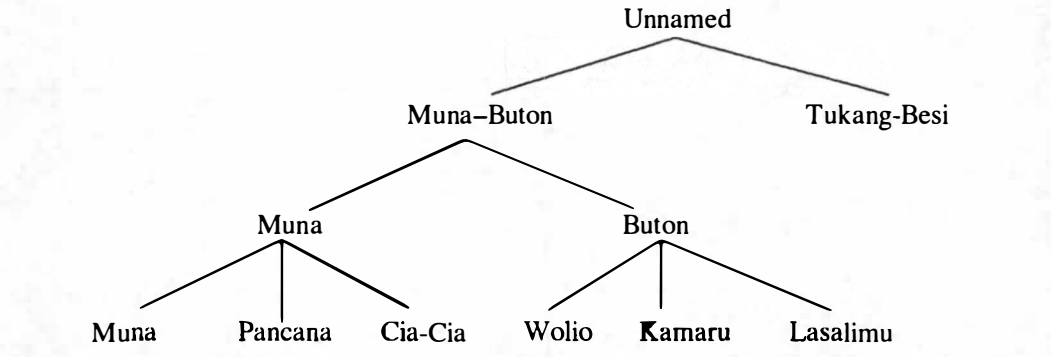



**Figure 3:** Anceau's (1978) subgrouping

The subgrouping claims made about the Muna–Buton group can be tabulated as in Table 1, where the same numeral in a column refers to an author placing all the languages bearing that notation in the same subgroup.

**Table 1:** The Muna–Buton group

| | Esser | Salzner | Anceaux | Bhurhanuddin | Kaseng | Grimes & Grimes | This paper |
|---|---|---|---|---|---|---|---|
| Wotu | — | — | — | — | — | 1 | 1 |
| Laiyolo | 1 | 1, 4 | — | 1 | — | 1 | 1 |
| Wolio | 1 | (2a) | 1 | 2 | 1 | 1 | 1 |
| Kamaru | (2) | (2b) | 1 | 3 | 1 | — | 1 |
| Tukang Besi | 3 | 3 | 2 | 1 | 2 | 1 | — |
| Muna | 4 | 1 | 3 | 4 | 3 | — | 2a |
| Kaimbulawa | (4) | (2a) | — | — | — | — | 2a |
| Pancana | (4) | (2a) | 3 | 4 | 4 | — | 2a |
| Cia-Cia | 2 | (2a, 2b) | 3 | 4 | 5 | — | 2b |
| Lasalimu | (2) | (2b) | 1 | 5 | 6 | — | 2b |
| Kumbewaha | (2) | (2b) | — | — | — | — | 2b |

## 3 The Wotu language

Wotu is grouped by Esser in his Toradja (=Kaili-Pamona[3]) group, but according to Noorduyn (1991a:144) '... he [Esser] changed his opinion ... and concluded that it belonged to the Buginese group'. Salzner (1960) followed this tack and classified Wotu in his Makassar–Bugis subgroup of South Sulawesi languages.

Mills (1975:604–612) implied that he considered Wotu to belong to the Toraja family rather than with the South Sulawesi languages when he wrote that '... on balance we find the points in common between Wo[tu] and Tor[aja] languages not only more numerous, but weightier' and discounted its putative connection with the South Sulawesi languages, later writing '... in my opinion, Wotu cannot claim a direct genetic affiliation with PSS [Proto South Sulawesi]'.

Finally, as mentioned above, in their lexicostatistic survey Grimes and Grimes (1987: 62–63) included Wotu in an extended Muna–Buton group, a move described by Noorduyn (1991a:144) as being 'The best solution to the problem ...' This solution is also favoured by Sirk (1988:11), who writes that 'What seems much more likely is that Wotu, Layolo and Wolio, possibly with some unknown dialects of Buton, etc., constitute a separate group which does not embrace Muna'.

A more detailed summary of the history of the debate surrounding the position of Wotu can be found in Noorduyn (1991b), but the main features of the hypotheses concerning the extent of Muna–Buton and the position of Wotu are summarised in Table 2:

**Table 2:** The subgrouping of Wotu

| | Esser | Salzner | Mills | Grimes & Grimes | Sirk (1988) | This paper |
|---|---|---|---|---|---|---|
| S. Sulawesi | 1 | 1 | 1 | 1 | 1 | 1 |
| Kaili-Pamona | 2 | 2 | 2 | 2 | 2 | 2a[4] |
| Wotu | 1, 2 | 1 | 2 | 3 | 3 | 2b |
| Laiyolo | 3a | 3a | 3 | 3 | 3 | 2b |
| Wolio | 3a | (3b) | — | 3 | 3 | 2b |
| Kamaru | (3b) | (3c) | — | — | (implied: 3) | 2b |
| Tukang Besi | 3c | 3d | — | 3 | — | — |
| Muna–Buton | 3b, d | 3b, 3c | — | 3 | 4 | 3a, b |

The presence of the same number in the entry for two languages in a column implies that the author in question considered the two languages to belong to the same grouping; Grimes and Grimes, for example, grouped Wotu, Laiyolo and Wolio together, but all apart from the South Sulawesi languages or the Kaili-Pamona languages. Numbers in different columns are not comparable. 'Tukang Besi' indicates both the languages of the Tukang Besi islands in Southeast Sulawesi, and the geographically remote Bonerate speech community. A dash (—)

3 The genetic unity of the Kaili-Pamona family has been demonstrated by Martens (1989).

4 Although not explicitly the subject of this paper, the Kaili-Pamona group (for references see Martens 1989), but some data are given in Table 1: Ledo, Napu, Pamona and Uma are all Kaili-Pamona languages) does share many sound changes with the Wotu-Wolio group, such as *e > *a*, *q > Ø, *R > Ø, *uy > *o* or *u*, and *Z > *d*.

shows that the language was not considered by the author, and a number in brackets shows that the area of the language would indicate its placement in that subgroup, even though it was not explicitly mentioned by that author.

It is the aim of this paper to present evidence supporting the conclusion that Wotu can be grouped with Laiyolo, Wolio and Kamaru, and that these languages can be better thought of as not belonging to the Muna–Buton group. The question of the internal relationships of the Muna–Buton group is not specifically addressed in this paper.

## 4 Issues

As can be seen from the brief summary presented, quite a lot of speculation has focused on the position of the Wotu language, and the genetic unity of the Muna–Buton group has not been challenged since Esser declared its existence. The assumption that all the languages of Muna, Buton and the Tukang Besi islands in Southeast Sulawesi, and the languages of southern Selayar (including the islands to the south) belong to one subgroup has only been questioned in any manner at all by Bhurhanuddin and by Sirk, and then only speculatively. The respective positions of Wolio and Wotu, and the question of which languages can be considered to be subgrouped with them, is therefore primary in an investigation of the extent of the 'Muna–Buton' group.

The questions addressed here are the following:

1a. Does Wolio show genetic unity with the other languages of Muna and Buton?

1b. If not, with what language(s) is Wolio affiliated?

2. What is the extent of the proposed language group that contains Wotu?

In answering these I present data from two previously known and two previously undescribed languages typical of the languages from the islands of Muna and southern Buton[5] — Muna, Cia-Cia, Kumbewaha[6] and Kaimbulawa[7] — and compare these with the languages of the putative Wotu–Wolio group that includes the Wolio, Kalao/Laiyolo,[8] and Wotu languages, as well as the previously undescribed Kamaru language of eastern Buton. Not addressed here is the question of the internal relationships of the subgroup that contains the remaining languages of the old Muna–Buton subgroup.

## 5 Approach and methods

I adopt the comparative method, specifically the examination of sound changes in the languages concerned, with the aim of adequately subgrouping through a body of shared innovations displayed by one or other of the groups.

Data were obtained from the following sources:

---

5 In addition to Muna along the west coast, the Kulisusu and Taloki languages of the Bungku language are found in the north of Buton island. These are not part of the Muna–Buton group, although clearly closely related, and are not considered here.

6 Located on the east coast of Buton, closely related to Lasalimu and Cia-Cia.

7 Spoken on the east coast of the island of Siompu off southwest Buton; closely related to Muna.

8 Dialects of the same language.

| | 'stone' *batu | 'hair' *buluq | 'flower' *buŋa | 'lip' *bibiR |
|---|---|---|---|---|
| Wotu | batu | bulu | buŋa | *sumba* |
| Laiyolo | batu | bulu | buŋa | *ŋinsu* |
| Kalao | batu | bulu | buŋa | *mɨnsu* |
| Wolio | batu | bulu | buŋa | biβi |
| Kamaru | atu | *potu* | buŋa | biβi |

Note that although all the languages lenite /b/ to *w* intervocalically, only the Muna–Buton languages do so initially. Muna also uses *kambea* for 'flower'.

The lenition of *b to *β* is not without its problems. Van den Berg (1991c:10–12) discusses the sound changes applying to certain etyma containing *b in seven Sulawesi languages, from the Kaili area in Central Sulawesi to the Muna–Buton area. Rearranging his table, and adding the languages relevant to this paper, produces the following chart of selected widespread etyma that reflect PAn *b, with non-lenited forms highlighted:

**Table 4:** Reflexes of PAn *b in languages of Central and Southern Sulawesi

| | Wol | K/L | Kam | Wot | Led | Kai | Mun | Kum | C-C | Nap | Pam | Uma |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *be(R)say | **b** | – | **b** | – | w | **b** | **b** | **b** | **b** | **b** | w | w |
| *beŋel | **b** | **b** | [p] | ? | **b** | **b** | **b** | [p] | – | **b** | w | w |
| *baqeRu | **b** | **b** | **b** | **b** | – | **b** | **b** | β | β | w | **b** | **b** |
| *benaqi | **b** | **b** | – | **b** | **b** | – | **b** | – | –† | (w) | **b** | w |
| *b-in-ahi | **b**/β | **b** | **b**/β | **b**/w | (m)**b** | -/**b** | -/**b** | β/β | β | –/β | – | **b** |
| *babuy | **b**/β | **b**/v | **b**/β | ? | **b** | β/β | β/β | β/β | β/β | – | w | w |
| *bulan | **b** | – | – | **b** | **b** | **b** | β | β | β | w | w | w |
| *bibiR | **b** | **b** | **b** | **b** | **b** | β | β | β | β | w | w | w |
| *buluq | **b** | **b** | **b** | **b** | **b** | β | β | β | β | w | w | w |
| *buŋa | **b** | **b** | **b** | **b** | **b** | – | β | – | **b**~β | **b** | – | w |
| *babaw | **b** | – | **b** | ? | **b** | β | β | β | β | – | w | – |
| *batu | **b** | **b** | **b** | **b** | β | β | β | β | β | w | w | w |

Key: Wol: Wolio; K/L: Kalao/Laiyolo; Kam: Kamaru; Led: Ledo; Wot: Wotu; Kai: Kaimbulawa; Mun: Muna; Nap: Napu; Kum: Kumbewaha; C-C: Cia-Cia; Pam: Pamona; Uma: Uma. Translations: paddle, deaf, new, sand, woman, pig, moon, lips, body hair, flower, above, stone (respectively). ? = not found in sources; – = no cognate in lists consulted; / separates twin reflexes in the same word; ~ shows alternation in different morphemes; Cia-Cia has *buŋa* 'flower' but *wunga* 'finger' (lit. 'flower-(of-hand)'). †Van den Berg cites *b* as the Cia-Cia reflex of *benaqi, but my lists all show *h* reflecting PAN *qenay. ( ) Napu *wungi* 'sand' is possibly not cognate with *benaqi' (van den Berg 1991c:12). [ ] unexpected devoicing in these two (geographically close) languages.

Table 4 again shows a clear split in reflexes between Wolio, Kalao, Laiyolo and Kamaru (and in most cases Ledo also), which consistently reflect *b as *b* (and as *ɓ*, in at least the cases of Wolio and Kamaru, for which I have collected the data myself) on the one hand and the other languages, which show various degrees of lenition in their treatment of *b. The reflexes of *be(R)say and *beŋel show particular resistance to this lenition in most languages outside Central Sulawesi, and *baqeRu, *binaqi and *b-in-ahi also show retention of *b as *b*

in the Munic languages (represented here by Muna and Kaimbulawa), and Central Sulawesi (excluding Napu). The reflexes for PAn *b-in-ahi 'woman' often show a reduplication of the first syllable, thus reflecting *(ba)-b-in-ahi, and the reflexes are given as for this form; a dash (–) shows that the language in question does not reflect the reduplicated form. The data in Table 1 are clear support for the idea that a sound change proceeds lexically through a language; the *b > *β* sound change is most advanced in Kumbewaha, but even there it has not spread to the reflex of *be(R)say. On the other hand, it appears that the sound change *b > *β* has just started in Ledo, beginning with the reflex of *batu.

**Table 5:** Reflexes of PAn *e
(see also 'new' under *q in Table 7)

| | 'three'<br>*telu | 'black'<br>*ma-qitem | 'six'<br>*enem |
|---|---|---|---|
| Kaimbulawa | totolu | mo/hito | nonoo |
| Muna | tolu | ɣito | noo |
| Kumbewaha | totolu | mo/kito | nonoʔo |
| Cia-Cia | totolu | mo/kito | nonoʔo |
| Wotu | talu/aŋo | ma/eta | ana |
| Laiyolo | talu | etta | ana |
| Kalao | talu | etta | anaŋ |
| Wolio | talu | ma/eta | ana |
| Kamaru | talu/aŋo | ma/eta | ana |

Here it is clear that whilst Kumbewaha, Kaimbulawa, Muna and Cia-Cia regularly reflect *e as *o*, the other languages show *a* reflexes.

**Table 6:** Reflexes of PAn *j

| | 'name'<br>*ŋajan | 'sun'<br>*qalejaw | 'gall'<br>*qapeju |
|---|---|---|---|
| Kaimbulawa | *kona* | holeo | — |
| Muna | nea | ɣoleo | ɣufei |
| Kumbewah | ŋea | holeo | — |
| Cia-Cia | ŋea | holeo | hopiu |
| Wotu | *saŋa* | mata/iyo | — |
| Laiyolo | *saŋa* | mataʔadžo | — |
| Kalao | *saŋa* | matanaajo | *pidu* |
| Wolio | *saro* | eo | — |
| Kamaru | *saro* | mata na/eu | *ma/pai* |

All the languages show *j > *y, which has affected the quality of the preceding vowel before disappearing. *Saŋa* in Wotu, Laiyolo and Kalao probably reflects a borrowing from a South Sulawesi language (compare Mandar, Mamuju, Toraja *saŋa*, Bugis *asiŋ*). Kalao and Kamaru forms for 'gall' are irregular; *pidu* probably being the result of borrowing, and *mapai* likely to reflect *ma-paqit 'bitter'.

**Table 7:** Reflexes of PAn *q
(see also 'snake' under *R and 'rain' under *Z)

| | 'new' *baqeRu | 'leg' *qaqay | 'white' *putiq | 'ten' *puluq |
|---|---|---|---|---|
| Kaimbulawa | bo̲ho̲u | ha̲he | mo/pute̲ | ompulu̲ |
| Muna | buɣou | ɣaɣe | pute̲ | ompulu̲ |
| Kumbewaha | βuk̲ou | k̲ake | mo/pute̲ | ompulu̲ |
| Cia-Cia | βuk̲ou | k̲ake | mo/pute̲ | ompulu̲ |
| Wotu | ku/ba | adže | ma/puti̲ | sapulu̲ |
| Laiyolo | bao | *biŋkiŋ* | putti̲ | spulu̲ |
| Kalao | *beʔeru* | *biŋkiŋ* | puti̲ | sapulu̲ |
| Wolio | b̲aau | ae | ma/puti̲ | sapulu̲ |
| Kamaru | b̲aau | ae | ma/puti̲ | sapulu̲aŋu |

PAn *q is preserved as *k/h* in Kumbewaha, Cia-Cia and Kaimbulawa, as *ɣ* in Muna, but dropped unconditionally in the other languages. Note the effect on a preceding *-i-* in *putiq. Kalao *beʔeru* 'new' is likely to be the result of the influence from a South Sulawesi language; compare with Makasar *beru*, Rongkong *baʔru*.

**Table 8:** Reflexes of PAn *R

| | 'blood' *DaRaq | 'thorn' *DuRi | 'snake' *qulaR | 'egg' *qateluR | 'wash' *DiRuq |
|---|---|---|---|---|---|
| Kaimbulawa | ʕe̲a | ki/ʕi̲ | hule̲ | hinteli̲ | *baho* |
| Muna | rea | ki/ri | ɣule̲ | ɣunteli̲ | ka/diu |
| Kumbewaha | xe̲a | xui | kule̲ | cikolu | *b̲aho* |
| Cia-Cia | re̲a | rui | *saʔa* | cikolu | *pib̲aho* |
| Wotu | raa | rui | ulo | *burau* | man/riyu |
| Laiyolo | raʔa | ruwi | ulo | *girau* | pɨn/riyu |
| Kalao | raʔa | rui | ulo | *korau* | pan/diu |
| Wolio | raa | rui | ulo | ontolu | *b̲aho* |
| Kamaru | raa | rui | ulo | ntolu | *petambusi* |

Note the non-phonemic glottal stop between like vowels in Laiyolo, Kalao *raʔa* 'blood'. PAn *R > *y in Kaimbulawa, Kumbewaha, Muna and Cia-Cia, which has affected the quality of the preceding vowel before disappearing. *R is lost without trace in the Woyu-Wolio languges. Note the divergent reflexes of **r ( < *D) in Kaimbulawa and Kumbewaha.

**Table 9:** Reflexes of PAn *uy

| | 'fire' *Sapuy | 'swim' *[l/n]aŋuy | 'pig' *babuy |
|---|---|---|---|
| Kaimbulawa | ipi̲ | leni̲ | βeβi̲ |
| Muna | ifi̲ | leni̲ | βeβi̲ |
| Kumbewaha | api̲ | leŋu | βeβi̲ |
| Cia-Cia | api̲ | pika/naŋu | βaβi̲ |

| | 'fire'<br>*Sapuy | 'swim'<br>*[l/n]aŋuy | 'pig'<br>*babuy |
|---|---|---|---|
| Wotu | apu̲ | mi/naŋo̲ | — |
| Laiyolo | apu̲ | pɨ/naŋo̲ | — |
| Kalao | apu̲ | pa/naŋo̲ | bavu̲ |
| Wolio | *ßaa* | po/ŋano̲ | baßu̲ |
| Kamaru | apu̲ | po/naŋu̲ | baßu̲ |

Apart from Kumbewaha and Cia-Cia *leŋu* and *pikanaŋu*, which irregularly reflect the *uy of *[n/l]aŋuy as *u*, the different reflexes are clearly split into two groups.

**Table 10:** Reflexes of PAn *w

| | 'eight'<br>*walu | 'nine'<br>*siwa | 'right'<br>*wanan |
|---|---|---|---|
| Kaimbulawa | o̲alu | si/siu̲a | s/u̲ana |
| Muna | o̲alu | siu̲a | s/u̲ana |
| Kumbewaha | o̲alu | so̲sia | mo/u̲ana |
| Cia-Cia | o̲alu | siu̲a | s/o̲ana |
| Wotu | walu | sasio | — |
| Laiyolo | falu | siyo | — |
| Kalao | valu | sio | kaana |
| Wolio | ßalu | sio | — |
| Kamaru | alu | sioaŋu | ka/antamo |

The *w is retained in all the languages, as a phonemic vowel in the Muna–Buton languages, and sporadically as a rounding of a following vowel or as a labial continuant amongst the Wotu-Wolio languages. See van den Berg (1991c:10–12) for a discussion of the problems associated with reflexes of *w in Muna and other Sulawesi languages.

**Table 11:** Reflexes of PAn *z/Z

| | 'chin'<br>*qaZay | 'path'<br>*Zalan | 'rain'<br>*quZan |
|---|---|---|---|
| Kaimbulawa | as̲e | s̲ala | his̲e |
| Muna | ɣas̲e | s̲ala† | ɣu̲se |
| Kumbewaha | as̲e | s̲ala | kia |
| Cia-Cia | hae | lala | kia |
| Wotu | — | d̲ala | ud̲a |
| Laiyolo | — | d̲ala | ud̲a |
| Kalao | ad̲e | d̲ala | ud̲a |
| Wolio | ad̲e | ʒ̲ala | *ßao* |
| Kamaru | ad̲e | ʒ̲ala | *monda* |

† *sala* is the South Muna form; North (standard) Muna has *kaŋkaha* for 'road'.

PAn *Z > *d* is clear in the Wotu-Wolio languages, whilst Kaimbulawa, Kumbewaha, Muna and Cia-Cia all show *Z > *s*. The loss of *s* in Kumbewaha *kia* 'rain' is unexplained. Kamaru *monda* is probably a borrowing from Tukang Besi *monda* 'rain'.

The sound changes relevant to the subgrouping of the Muna–Buton and Wotu–Wolio languages exemplified in the data sets above are summarised in Table 12.

**Table 12:** Muna–Buton vs Wotu-Woliu sound changes

| *PAn | *b | *e | *j | *q | *-iq# | *R | *uy | *w | *Z |
|---|---|---|---|---|---|---|---|---|---|
| Kaimbulawa | β | o | *y | h | e | *y | i | o,u | s |
| Muna | β | o | *y | ɣ | e | *y | i | o,u | s |
| Kumbewaha | β | o | *y | k | e | *y | i | o,u | s |
| Cia-Cia | β | o | *y | k, h | e | *y | i | o,u | s |
| Wotu | b | a | *y | Ø | i | Ø | o, u | __o, w† | d |
| Laiyolo | b | a | dž | Ø | i | Ø | o, u | __o, f | d |
| Kalao | b | a | j | Ø | i | Ø | o, u | __Ø, o, v | d |
| Wolio | b | a | *y | Ø | i | Ø | o, u | __o, β | d |
| Kamaru | b | a | *y | Ø | i | Ø | u | __Ø, o | d |

† A PMP *w is not always directly reflected in Wotu, Laiyolo, Kalao, Wolio or Kamaru, but is apparent in the rounding of the following vowel, such as Wolio *sio* 'nine', where the *o* reflects the rounding of the original *a under the influence of the *w*.

## 7 The Wotu–Woliu languages

There are two languages on Buton that do not conform to the expected patterns of sound changes in that area, Wolio (the language of the Sultanate capital in Baubau) and Kamaru, a small language in eastern Buton. When these languages are compared with Laiyolo/Kalao and Wotu from South Sulawesi, their relationship with the other Muna–Buton languages can be seen to be significantly less close than with the languages exained from South Sulawesi.

The emergence of two different subgroups in the remaining languages of Muna and Buton is obvious, with Kaimbulawa, Muna, Kumbewaha and Cia-Cia forming one group, and the Wotu, Laiyolo, Kalao, Wolio and Kamaru languages forming the second. Convincing sound changes outlining these two groups are the treatment of PAn *b, *e, *-iq#, *R, *uy, and *z/Z. The effect of a final *q on a preceding vowel is interesting, in the light of Sirk's (1989:57) comment that 'The lowering of high vowels before final -q unites the SSul [South Sulawesi] languages' (as against Kaili-Pamona, Wolio, Laiyolo and Wotu). The data presented here have validated this statement with respect to the languages under question, and shown that in the Muna–Buton languages, while a *u* is unaffected by a final *-q, the sequence *-iq lowers to *-e*, different to the pattern found in the Wotu-Wolio languages, in which both high vowels are unaffected by a following *q. The proposed subgrouping for these languages is as shown in Figure 4.

Not all the languages in Figure 4 have been discussed in this paper, nor all the evidence used to subgroup them, such as the presence of verb classes in the languages of the Munan subgroup, or the innovations found in different groups in terms of pronominal indexing on the verb; these remain as topics for a later, more detailed discussion of the Muna–Buton languages.

The data show that the Wolio language does not belong to a sensible low-level subgrouping that includes these other languages of Muna and Buton, apart from its relationship with Kamaru. These two languages are at the end of a very spread-out language family ranging from Wotu at the northern end of the Gulf of Bone, down to Kalao island in the Sea of Flores, and up to Buton. This family of languages cannot be closely linked with the other (presumed original) languages of Muna and Buton. In support of the claim that Wolio and Kamaru are not indigenous to the area, Bhurhanuddin reports that the Wolio tradition is that they are immigrants to the area;[10] the name of the sultanate capital city, Baubau (< Wolio *baau-baau* < *baau* 'new'), also supports this claim.[11]

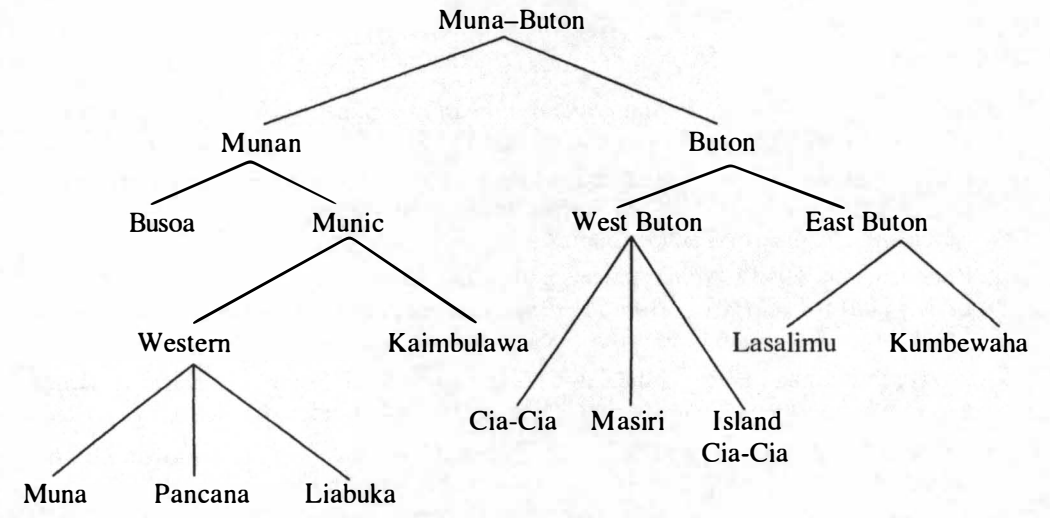



**Figure 4:** Proposed subgrouping

On the other hand, the evidence linking Wotu to Wolio is compelling. Friberg and Laskowske (1989:14) report that Kalao is said to be the original language of its area, writing '... in deference to folk history which makes Kalao original, followed by Barang-barang and Laiyolo, in turn followed by Wotu', and Wolio people in Ujung Pandang have reported to me the belief that the ancestors of the Wolio people came from the western side of the Gulf of Bone, affirming the belief that they are not indigenous to Southeast Sulawesi. The speakers of the Kamaru language in eastern Buton acknowledge that their origin is in the Wolio area, making them a later movement from the Wolio area to the east of the island. Despite the evidence that these languages do subgroup together, we cannot at this stage propose subgrouping within the family.

10 Bhurhanuddin (1979:48): 'Tradisi Wolio menang mengungkapan bahwa nenek moyang mereka adalah pendatang di Buton'.

11 Mike Southon (pers. comm.) reports that members of the palace nobility in Baubau are proud that they are immigrants to the area, in contradiction to the popular sentiment.

## 8 Conclusions

While the unity of the Wotu-Wolio group has been demonstrated, reducing the size of the old Muna–Buton group, the extent of the languages in the new Muna–Buton group remains to be determined, though the evidence points to there being two subgroups (see fn.8). Subgrouping within the Wotu-Wolio group, and the question of the affiliations of this group in the larger Sulawesi linguistic picture, remain to be addressed. A genetic relationship between all the languages discussed here at a higher level is likely in view of the sound changes that they do share, such as the loss of final consonants, and the development of PAn *D > **r,[12] and the treatment of *w and *j, but a detailed examination of this hypothesis is outside the scope of this paper.

## References

Anceaux, J.C., 1952, *The Wolio language: outline of grammatical description and texts.* KITLV, Verhandelingen 11 (second edition 1988). 's-Gravenhage: Martinus Nijhoff.

—— 1978, The linguistic position of South-East Sulawesi. In Stephen Wurm and Lois Carrington, eds *Papers from the Second International Conference on Austronesian Linguistics.* Canberra: Pacific Linguistics.

Berg, René van den, 1991a, Muna dialects and Munic languages: towards a reconstruction. In Ray Harlow, ed. *Papers from the Fifth International Conference on Austronesian Linguistics.* Auckland: Linguistics Society of New Zealand.

—— 1991b, Preliminary notes on the Cia-Cia language (South Buton). In Harry A. Poeze and Pim Schoorl, eds *Excursies in Celebes*, 305–324. Leiden: KITLV Uitgeverij.

—— 1991c, Muna historical phonology. In J.N. Sneddon, ed. *Studies in Sulawesi Liguistics part 2.* NUSA vol. 33:1–28. Jakarta: Universitas Katolik Indonesia.

Bhurhanuddin, B.H., 1979, *Bahasa-bahasa Daerah di Sulawesi Tenggara.* Kendari, MS.

Donohue, Mark, n.d., Field notes in the languages of Southeast Sulawesi.

Esser, S.J., 1938, Talen. In *Atlas van tropisch Nederland.* Amsterdam: Koninklijk Nederlandsch Aardrijkskundig Genootschap: blad 9b.

Friberg, T., 1987, Unpublished Kalao word list.

Friberg, T. and T.V. Laskowske, 1989, South Sulawesi languages. In Sneddon, ed. 1989:1–17.

Grimes, Charles E. and Barbara D. Grimes, 1987, *Languages of South Sulawesi.* Canberra: Pacific Linguistics.

Kaseng, Syahrudin, Alimuddin D.P., Andi Mahmuddin and Rasdiana P., 1987, *Pemetaan Bahasa-bahasa di Sulawesi Tenggara.* Jakarta: Departemen Pendidikan dan Kebudayaan.

Martens, Michael P., 1989, *Proto Kaili-Pamona: reconstruction of the protolanguage of a language subgroup in Sulawesi.* SIL, Sulawesi: MS.

Mills, Roger F., 1975, *Proto South Sulawesi and Proto Austronesian phonology.* PhD thesis, University of Michigan.

Noorduyn, J., 1991a, The languages of Sulawesi. In H. Steinhauer, ed. *Papers in Austronesian Linguistics* No.1: 137–150. Canberra: Pacific Linguistics.

---

12 With certain exceptions, such as Muna *ka/diu* < *DiRuq 'wash'. The majority of cases, however, do illustrate a *D > *r* correspondence.

—— 1991b, *A critical survey of studies on the languages of Sulawesi.* KITLV Bibliographical series 18. Leiden: KITLV Press.

Salombe, C. et al., 1987, *Struktur Bahasa Wotu.* Jakarta: Departemen Pendidikan dan Kebudayaan.

Salzner, Richard, 1960, *Sprachenatlas des indopazifischen Raumes.* Map 23–D (Makassar – Butung). Wiesbaden: Otto Harrassowitz.

Sirk, Yu. Kh, 1988, Research needs: the Wotu language. In *Baruga: Sulawesi Research Bulletin* 2:10–12

Sirk, Ü., 1989, On the evidential basis for the South Sulawesi language group. In Sneddon, ed. 1989:1–17.

Sneddon, J., ed., 1989, *Studies in Sulawesi linguistics* part 1. NUSA vol. 31. Jakarta: Universitas Katolik Indonesia.

# 4 *Lexical similarity, sound change and intelligibility of Atayalic dialects*

DER-HWA VICTORIA RAU

## 1 Introduction[1]

The Atayalic language group comprises two major subgroups: Atayal and Sediq. The former can be further divided into two major dialects: Squliq and C'uli'. In terms of the degree of dialectal divergence, C'uli' dialects are considered to be the most divergent, followed by Sediq, while Squliq dialects are fairly uniform (Li 1981).

Li (1980, 1981, 1982a,b, 1985, 1996) has contributed substantially to our understanding of the classification and phonology of the Atayalic groups. Other phonological studies include Squliq Atayal (Egerod 1966; Hirano 1972; Yamada & Liao 1974; Chiang 1996) and Sediq (Yang 1976). Detailed phonological accounts of individual C'uli' varieties are, however, lacking. This study is an attempt to further clarify the relationship between the three Atayalic dialects of Ren-Ai Township, Nantou County.

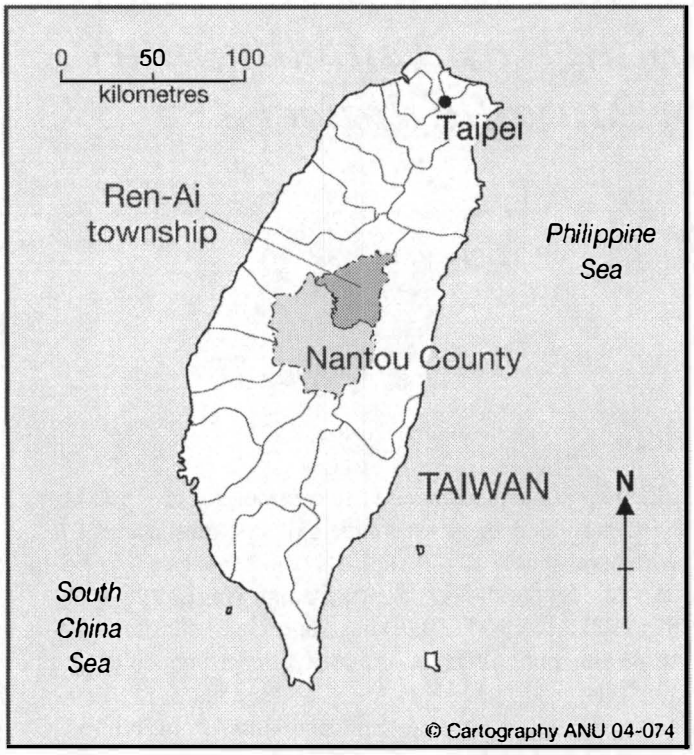
Ren-Ai Township is located in central Taiwan, as illustrated in Map 1, and is claimed to be the 'Atayalic homeland' (Li 1993) due to its great dialectal diversity. As shown in Map 2, except for two Bunun villages (1. Zhong-Zheng, 2. Fa-Zhi) to the southwest of Nantou County and two Han villages (14. Rong-Xing, 9. Da-Tong) toward the northeast, the rest of the county houses mostly speakers of Atayalic dialects with Atayal dialects on the north and Sediq dialects on the west and the south. Wan-Da, in Qin-Ai village (4), is the only C'uli' Atayal speech community, surrounded by Sediq speakers (Map 2).

There are different reference terms used in the paper. Under the township, there are Chinese village names, followed by the Chinese names of the speech community. Mstbaun, Inago, and Palngawan are the names Atayalic people use to refer to their own groups and have been adopted by Li (1980, 1982a) in his description of those groups and their dialects. The distribution of the local Atayalic dialects in the township is illustrated in Map 3. The

1 This study was supported by a grant from the National Science Council for the project 'Sociolinguistic Survey of the Atayalic Dialects of Nantou County', 8/1/1997–7/31/1998 (NSC-87-2411-H-126-010).

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 37–95.
Canberra: Pacific Linguistics, 2004.

local Atayalic dialects belong to three major dialects, as classified by Li. Mstbaun is a Squliq dialect. B'ala' and Palngawan are C'uli' dialects. Tongan, Toda, Truwan, and Inago are Sediq dialects.

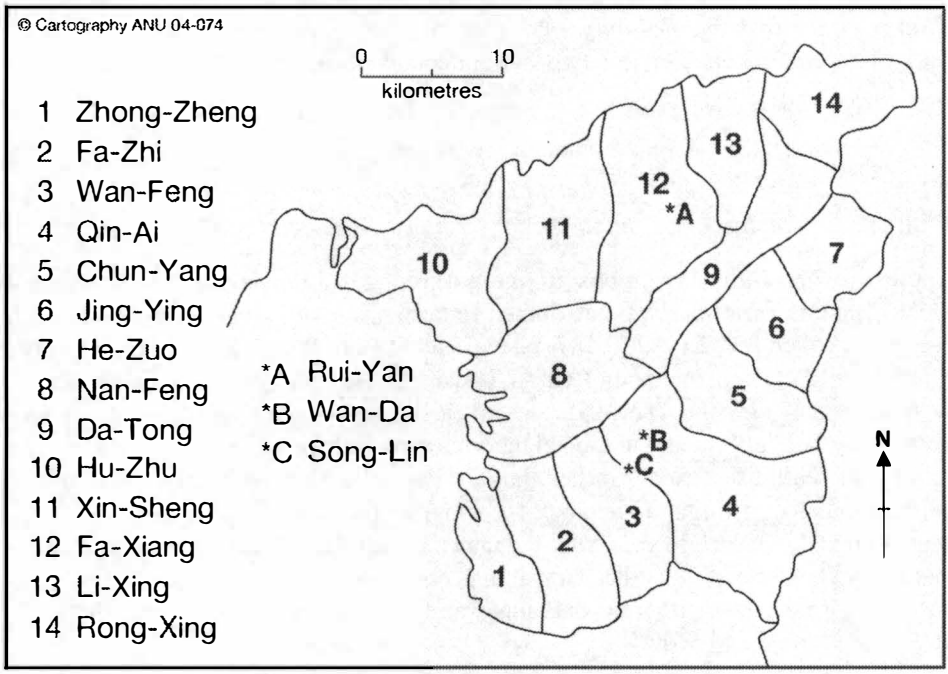



**Map 1:** Ren-Ai Township, Nantou County in Taiwan

Three speech communities, Rui-Yan (Fa-Xiang Village), Wan-Da (Qin-Ai Village), and Song-Lin (Qin-Ai Village), representing the three dialects, Squliq, C'uli', and Sediq respectively, were chosen as reference sites for the study. Wan-Da and Song-Lin communities are within walking distance of each other while linguistically Palngawan and Inago are classified as Atayal and Sediq respectively. Language contact between the two communities is inevitable. Rui-Yan, on the other hand, is geographically separated from the other two communities although linguistically Mstbaun is classified as an Atayal dialect, more closely related to Palngawan than Inago.

The total population of the fourteen villages of Ren-Ai Township was 15,143 as of 1996. 76% of the residents are aborigines. Over 90% of the population of the two villages in our study are Atayalic.

The presentation of this paper is organised as follows. After this introduction, a review of Li's studies of the dialect situation is presented in §2. Section 3 introduces my data, followed by the cognate percentages in §4 and sound correspondences in §5. Section 6 summarises the findings of the two previous sections. Section 7 discusses the results from dialect intelligibility testing. The Appendices contain the following kinds of data: Appendix 1. Word lists, part 1: 326 lexical items, part 2: 190 lexical items, Appendix 2. Recorded text tests for intelligibility.




**Map 2:** Villages in Ren-Ai Township of Nantou County

## 2 Atayalic dialects

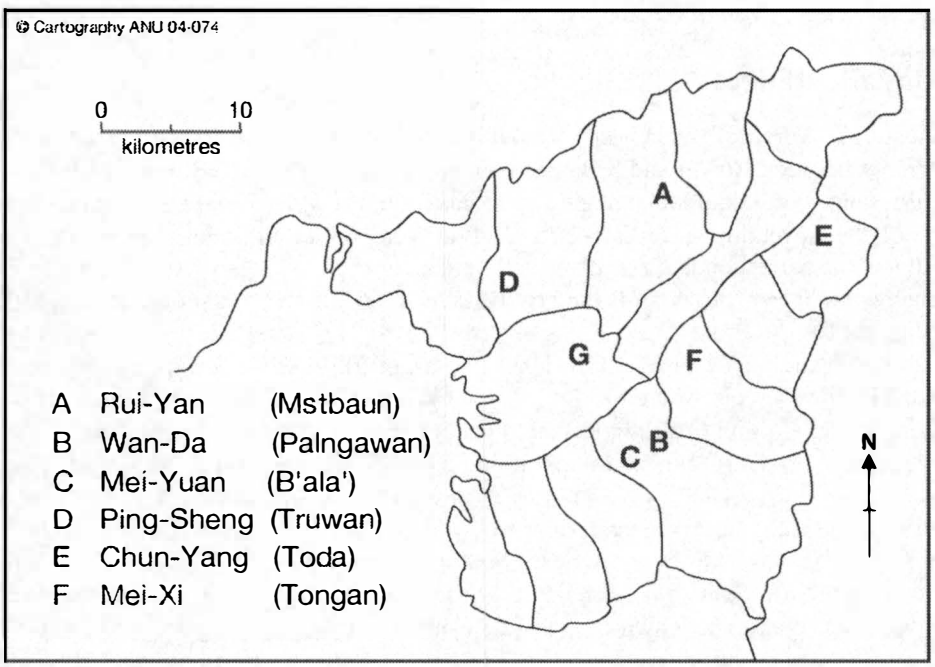
The term 'Atayalic' refers to both Atayal and Sediq. The important phonological differences between Atayal and Sediq, as pointed out by Li (1980, 1985), are as follows: (1) Atayal retains word-final labial stops and nasals /p, b, m/ while Sediq has changed to velars /k, ŋ/; (2) Sediq retains voiced stops /b, d, g/ in word-initial and medial positions whereas Atayal has the corresponding fricatives and liquids /β, r, ɣ/; (3) Sediq retains /r/ while Atayal has changed to /y/, /z/ or zero; (4) the Proto Atayalic *-d has reflexes *-t* or *-ʔ* in Atayal, but *-c* in Sediq; (5) for the Proto Atayalic *-g-, Atayal has *-g-* [ɣ] as reflexes, while Sediq has *-r-*; (6) Proto Atayalic *-g'- has Atayal reflexes *-r-*, *-s-*, or trill *r*, while Sediq generally has *-y-* if preceded by /i/ or *-g-* elsewhere. Li (1980, 1996) also presented Tsuchida's three criteria for subgrouping Squliq and C'uli' dialects: (1) phonological, (2) morphological, and (3) lexical differences. First, in terms of phonological differences, three types of phonetic correspondences were cited: (a) Squliq /s/ corresponds to C'uli' /c/ (< PA *c) as in /sbiŋ/ vs /cbiŋ/ 'sweet'; (b) Squliq /-r-/ corresponds to C'uli' /-s-/ (<PA *g') as in /pira'/ vs /pisa'/ 'how many', /kira'/ vs /kisa'/ 'a little later'; (c) Squliq /-ʔ/ corresponds to C'uli' /-t/ or /-c/ (< PA *-d) as in /qoliʔ/ vs /qolit/ 'rat'. Li (1996:188) presented the regular sound changes in the C'uli' varieties of Nan-Ao Township, Yi-Lan County as follows: (1) *q* > *ʔ* or *-ʔ-*, (2) *-p* > *-k*, *-m* > *-ŋ*, (3) loss of initial consonant, (4) /g-/ > *x*. Second, morphological differences, specifically pronominal differences, are illustrated with the following examples in Li (1980), Squliq /saku'/ or /ku'/ vs C'uli' /cu/, /ci/, /su/ or /si/ 'I'. Li (1996:188) generalised two types of morphological differences between Squliq and C'uli' varieties in Yi-lan due to the innovation of male forms: (1) last syllable or the last vowel/consonant differences, for

example *sas-aw* > *sas-iq* 'shade',[2] *bga-yaw* > *bga-ti?* 'Alocasia'; (2) insertion of an infix, e.g. *qmalup* > *qmalu-ya-k* 'hunt', *luhuŋ* > *luh-i-uŋ* 'mortar', *guquh* > *guq-il-uh* 'banana'. Third, lexical differences between the two subdialects include, to name a few, cited by Li (1996:185),

| | 'chicken' | 'shoulder' | 'sweat' | 'plant' | 'hit' |
|---|---|---|---|---|---|
| Squliq: | *ŋta?* | *qhiyaŋ* | *yabux* | *muya?* | *mihiy* |
| C'uli': | *wayluŋ* | *hŋali?* | *rinaŋ* | *muhi?* | *mahiy* |

There are certainly exceptions to these three general criteria, as indicated by Li (1996:185). The variations were attributed to borrowing and language contact and further research was called for. Li (1996:187) cited Tsuchida's (1980) finding of reflex *s* instead of *r* in many C'uli' dialects including Ren-Ai Township, Nantou County, for example *pgyaran* 'escape' instead of *pgyasan*. This indicates an early borrowing before dialectal diffusion.

Palngawan, a C'uli' dialect in the neighborhood of Sediq dialects, was found to share the phonological features of Sediq rather than Atayal. Li, therefore, concluded that lexical evidence is more useful than phonology for subgrouping Atayal and Sediq. Li (1985) later presented lexical evidence to show that Palngawan is an Atayal dialect because it shares 214 lexical items exclusively with other Atayal dialects, whereas only 11 lexical items are shared exclusively with Sediq. Furthermore, Palngawan has 50 unique lexical items, different from other Atayal dialects and Sediq.




**Map 3:** Atayalic speech communities in Ren-Ai County

2 The difference between Squliq /sasaw/ and C'uli' /sasi'/ 'shade' was classified as a lexical difference in Li (1980a) but was reclassified as a morphological difference in Li (1996).

Li (1980, 1982a) further proposed that the Atayalic group shares the same direction of sound change: (1) *-l* > *-n*; (2) *-t*, *-d* (only in Sediq) > *-c*; (3) *-b* > *-p*; (4) *-p* > *-k*, *-m* > *-ŋ*; (5) *-g* > *-w*, *-y*; (6) *c* > *s*, (7) vowel-deletion before stress. Age and gender both affect sound change, but the former was claimed to be more important than the latter. In the case of sound change from labials to velars in (4), Li reported his observation of the change in Palngawan dialect and cited Tsuchida's observation that the same change is completed in Mstbaun dialect. But our word-list elicitation of Mstbaun dialect shows that the sound change from labials to velars is still in progress. This finding will be discussed in §5. Li (1982a) certainly made several important observations and hypotheses on sound change, which await further quantitative analyses based on a methodology of sociolinguistic variation.

Li's studies (1980, 1982b, 1983) indicate that Mayrinax and Pa'nakuali' are the only two dialects of Atayal that show certain well-defined differences between the male and female forms of speech. The female forms preserve archaic features, whereas the male forms are innovative. But the majority of the male forms in Mayrinax are the ones currently used in other Atayal dialects and used as representative forms in the Atayalic word list in Li (1996).

Several phonemic and phonetic features in C'uli' which are different from Squliq were mentioned in Li (1980a) and are summarised as follows. In terms of phonemic differences, /e, o/ seem to be phonemic and /q/ does not occur in Maspazi'. Skikun and Mayrinax do not have /z/. As for phonetic differences, the bilabial fricative [β] is replaced by labiodental [v] in the speech of younger speakers of Maspazi'. The liquid [r] is commonly a flap in Maspazi', but is a retroflexed fricative [ɻ] or [ʐ] in Skikun. The devoicing of [ɣ] is completed in younger and female speakers of Skikun dialect. The palatalisation of /t/ before /i/ does not occur in Maspazi' and Mayrinax. Final /r/ occurs in Mayrinax and Palngawan. Vowels before penult and diphthongs are preserved in Maspazi' and Mayrinax.

Finally the sound systems of the three dialects in this study, based on Li (1980a, 1981) are summarised in Tables 1–3.

**Table 1:** Sound system of Mstbaun (Squliq)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | t | (c) | k | q | ʔ | | i | | | u |
| b[β] | | z | g[ɣ] | | | | | e | o | |
| | | s | x | | h | | | | a | |
| | l | | | | | | | | | |
| | r | | | | | | | | | |
| m | n | | ŋ | | | | | | | |
| w | | y | | | | | | | | |

**Table 2:** Sound system of Palngawan (C'uli')

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| p | t | c | k | | ʔ | | i | | u |
| b | | | g | | | | | | |
| | | s | x | | h | | | e | o |
| | l | | | | | | | | a |
| | r | | | | | | | | |
| | r̃ | | | | | | | | |
| m | n | | ŋ | | | | | | |
| w | | y | | | | | | | |

**Table 3:** Sound system of Inago (Sediq)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| p | t | (c) | k | q | ʔ | | i | | u |
| b | d | | g | | | | | ə | o |
| | | s | x | | h | | | a | |
| | l | | | | | | | | |
| | r | | | | | | | | |
| m | n | | ŋ | | | | | | |
| w | | y | | | | | | | |

## 3 Goals of the study

This study is a sociolinguistic survey of three Atayalic dialects of Ren-Ai Township, Nantou County, and a systematic comparison of Atayalic dialects to further clarify the subgroupings of the Squliq and C'uli' dialects of the Atayal and the Sediq language. The goals of the study are to investigate lexical similarity, dialect intelligibility and systematic sound change of the three dialects in Nantou. Three speech communities, Rui-Yan, Wan-Da, and Song-Lin, representing the three dialects Mstbaun Squliq, Palngawan C'uli', and Inago Sediq respectively, were chosen as reference sites for the study.

Since Ren-Ai Township is considered the 'homeland' of the Atayalic people and Rui-Yan is located in their 'place of origin', Mstbaun's status as a representation of Squliq dialect is certainly justifiable. Although Squliq and Palngawan are considered the most innovative and not comparable with Mayrinax in terms of value for historical reconstruction, they are nonetheless included as evidence for Li's reconstruction of Proto Atayalic phonology. Since Palngawan and Inago speech communities are within walking distance of each other while linguistically classified as Atayal and Sediq respectively based on Li's lexical evidence, a better understanding of Palngawan is important to clarify its relationship with other Atayalic dialects.

### 3.1 Research questions

Our quest is further divided into the following three questions:

(1) Is Palngawan more similar to Atayal or Sediq in terms of lexical evidence?

(2) Does Palngawan share more phonological features (i.e. sound change) with Atayal or Sediq?

(3) What are the levels of intelligibility among the three dialects?

### 3.2 Data

Two sets of word lists were used for analysis of lexical similarity and sound correspondences. One is the 326 Atayalic lexical items in Li (1981) with an addition of our Mstbaun data and a revision of Li's Palngawan and Inago data, the other is the 190 Atayal lexical items[3] (Li 1996:196–213) with addition of our Palngawan and Mstbaun data. All the data on Mstbaun, Palngawan, and Inago were collected by the author while the others are Li's. The word list with 326 items is a comparison among all three Atayalic dialects, while

[3] Li's (1996) word list contains 190 entries but the last four are sentences.

the list with 190 items is mostly concerned with the comparison between Squliq and C'uli'. Thus the first word list contains two Squliq dialects, four C'uli' dialects, and four Sediq dialects, while the second word list contains ten Squliq varieties and eight C'uli' varieties. The word lists of Mstbaun, Palngawan, and Inago that we collected are included in Appendix 1 along with the background information on our informants. All the word lists have reached a reliability code of C and above: 'average survey situation with good bilingual informants and satisfactory opportunity to double check' (Wimbish 1989:31). The shorter list with 190 lexical items was used with our older informants. There are only 74 overlapping items in both lists, cross-referenced in Appendix 1, part 1.

Three texts were recorded for the intelligibility test. A narrative text of personal experience approximately three minutes long was elicited from each reference site to be made into recorded text tapes for listening comprehension. Ten content questions for each text were abstracted from the story, translated into the three dialects, and dubbed onto the tapes as the test questions. The three texts for the RTT are included in Appendix 2.

## 4 Lexical similarity

Lexical similarity is usually cited to answer the question whether Palngawan is more similar to Atayal or Sediq. Based on Li's (1985) lexical evidence, Palngawan is found to share more exclusive lexical items with Atayal than with Sediq. However, the problems with Li's study are twofold: first, several different Atayal and Sediq dialects, ranging from Mayrinax to Squliq in Atayal and from Tongan to Inago in Sediq, are compared with Palngawan depending on which data are available for comparison; second, no criteria are given to determine lexical similarity, for example high strength of correspondence sets, so that there is no reason to believe that the word 'juice' *bu?* in Palngawan shares exclusively with *buq* in Mayrinax but not with *beyuq* in Tongan (1985:702). In our study, the languages for comparison were constant and a principle of quantification was established to group cognates.

### 4.1 WordSurv

Two sets of word lists were entered into the WordSurv computer program (Wimbish 1989) for analysis. After the word lists were entered into the computer with the cognate decisions made by the researcher, the program provided the following three types of information, which served as the basis to answer our first question: (1) shared vocabulary counting, (2) phonostatistic analysis of cognates, and (3) the COMPASS analysis to measure the strength of proposed phoneme correspondences and give an indication of the likelihood that words grouped in cognate sets are actually cognates.

#### *4.1.1 Shared vocabulary counting*

The 'shared' function of the Wordsurv program was used to produce the number of shared cognates as a percentage of the basic vocabulary. Since the first classification of the cognates was based on their appearance (apparent cognates) and the accurate determination of cognates depends on application of comparative method, these preliminary counts are used only for comparison.

### *4.1.2 Phonostatistic analysis*

The 'degrees of difference' (DD) analysis rather than the 'sound changes' analysis was adopted for the phonostatistic analysis. The degrees of difference between sounds is the number of minimal steps that would be required to change one to the other. Under this approach, all identical correspondences are counted as having a DD value of 0, while each nonidentical correspondence pair has a DD value of 1, regardless of its features. The advantage of this default strategy is that it avoids researcher bias in entering DD values into the computer, but it has the disadvantage of assigning small values to potentially large sound changes. Since the results of the 'shared' function of the program are used for comparison only, the default strategy is sufficient for our purpose.

### *4.1.3 The COMPASS analysis*

COMPASS, for 'Comparativist's Assistant', is an algorithm that was cited by Wimbish (1989:67) as having been developed by Donald Frantz (1970) based on the comparative method for linguists to determine genetic relationship between languages and to reconstruct the protolanguage. It is used to measure the likelihood that forms entered as cognates in the word-list database are in fact historically cognates, and does this by examining their frequency of occurrence in the data. The COMPASS algorithm was used to generate the following three tables: (1) phoneme correspondences, (2) item pairs list with cognate strengths, and (3) the number of word pairs within given ranges of strength. The strength index representing the likelihood that the correspondence is the result of a regular sound change was assigned by using the default threshold values of the program (upper threshold 15, lower threshold 2, bottom threshold 1). A correspondence with 15 or more occurrences scores a maximum strength of +1, representing the maximum confidence that it is regular correspondence. A correspondence with only 1 occurrence scores a maximum negative strength of −1, representing maximum confidence that it is not. A correspondence with 2 occurrences scores a medium negative strength of −0.5, while correspondences with between 3 and 14 ocurrences score a positive strength between 0 and 1 which grows proportionately with the number of occurrences. Values between the two extremes (+1 and −1) represent intermediate degrees of likelihood.

The pairs of correspondences that have the highest average segment strength of 1.00 in the COMPASS tabulation were chosen to represent true cognates. Other cognates with strengths greater than 0.85 but smaller than 1.0 are also compared.

## 4.2 Results of lexical similarity

The following section discusses the results of lexical similarity of the three Atayalic dialects based on a word-list analysis to answer the question whether Palngawan is more similar to Atayal or Sediq based on lexical evidence. We begin by examining the shared vocabulary counts of the Atayalic dialects and proceed to compare the cognates among the three dialects, obtained from the COMPASS analysis, then explain the differences between Li's (1985) results and ours.

### *4.2.1 Shared vocabulary counting*

Table 4 was generated using the 'shared' function of the WordSurv program to calculate the similarity percentages among Atayalic dialects. The set of word lists (326 words) comprises one Squliq dialect (Squliq in Taoyuan County), three C'uli' dialects (Maspazi? in Hsinchu County, Skikun in Yilan County, Mayrinax in Miaoli County), and three Sediq dialects (Tongan, Toda, Truwan) in Nantou County from Li's (1981) data and one Squliq (Mstbaun), one C'uli' (Palngawan), and one Sediq (Inago) in Nantou County from the current study (Appendix 1, part 1). The similarity percentages report the number of shared cognates as a percentage of the basic vocabulary compared.

In Table 4, the first two varieties from the top (Mstbaun and Squliq) are classified as Squliq dialects, the next four (from Palngawan to Mayrinax) as C'uli', and the last four (from Inago to Truwan) as Sediq. A first look at the similarity percentages between Palngawan–Mstbaun and Palngawan–Inago seems to indicate that Palngawan is slightly more similar to Mstbaun than Inago (85% vs 83%). But we are still far from being able to draw the conclusion that Palngawan is more similar to Atayal than Sediq because of the following results: (1) Mstbaun is more similar to Inago than Palngawan (86% vs 85%), (2) Squliq is more similar to Inago than Palngawan (92% vs 90%). These differences are so slight that they may not be significant. Squliq dialects also seem to be more similar to Sediq than to C'uli' dialects.

**Table 4:** Similarity percentages among Atayalic dialects

| Mstbaun | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 96 | Squliq | | | | | | | | | |
| 85 | 90 | Palngawan | | | | | | | | |
| 91 | 96 | 91 | Maspazi? | | | | | | | |
| 94 | 97 | 88 | 96 | Skikun | | | | | | |
| 90 | 96 | 90 | 97 | 97 | Mayrinax | | | | | |
| 86 | 92 | 83 | 89 | 89 | 88 | Inago | | | | |
| 90 | 98 | 90 | 95 | 95 | 95 | 96 | Tongan | | | |
| 91 | 98 | 89 | 96 | 95 | 96 | 96 | 100 | Toda | | |
| 91 | 98 | 89 | 95 | 95 | 96 | 97 | 100 | 100 | Truwan | |

Squliq dialects

C'uli' dialects

Sediq dialects

Tables 5 and 6 were generated the same way, but based on a different set of word lists including mainly dialects of Atayal. Table 5 compares Palngawan with other Squliq dialects while Table 6 compares Palngawan with other C'uli' dialects and Mstbaun with other Culi' dialects. The word lists (190 words) contain data from nine Squliq dialects (Pyanan, Lmuan, Habun Bazinuq, Syanuh, Kulu, ŋŋupa, Haga-Paris, Kubaboo, Rghayuŋ) and seven C'uli' dialects (Mnibu?, Mnawyan, Mkgugut, Pyahaw, Ryuhiŋ, Mtlaŋan, Knŋyan) in Yilan County from Li's data (1996) and one Squliq (Mstbaun) and one C'uli' (Palngawan) in Nantou County from the current study (Appendix 1, part 2).

**Table 5:** Similarity percentages between Palngawan and other Squliq dialects

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Mstbaun | | | | | | | | | | |
| 92 | Pyanan | | | | | | | | | |
| 88 | 91 | Lmuan | | | | | | | | |
| 94 | 91 | 91 | Habun Bazinuq | | | | | | | |
| 93 | 93 | 92 | 97 | Syanuh | | | | | | |
| 90 | 92 | 88 | 91 | 90 | Kulu | | | | | |
| 84 | 88 | 80 | 84 | 81 | 90 | Ŋŋupa | | | | |
| 91 | 92 | 88 | 88 | 90 | 97 | 88 | Haga-Paris | | | |
| 90 | 89 | 86 | 90 | 89 | 94 | 92 | 94 | Kubaboo | | |
| 92 | 91 | 88 | 92 | 92 | 95 | 90 | 95 | 98 | Rghayuŋ | |
| 66 | 65 | 65 | 66 | 64 | 68 | 70 | 67 | 67 | 68 | Palngawan[4] |

**Table 6:** Similarity percentages between Palngawan and other C'uli' dialects

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Palngawan | | | | | | | | |
| 62 | Mnibu? | | | | | | | |
| 60 | 96 | Mnawyan | | | | | | |
| 69 | 69 | 71 | Mkgugut | | | | | |
| 67 | 67 | 68 | 95 | Pyahaw | | | | |
| 66 | 68 | 69 | 96 | 94 | Ryuhiŋ | | | |
| 66 | 66 | 67 | 93 | 93 | 96 | Mtlaŋan | | |
| 68 | 65 | 67 | 94 | 93 | 94 | 93 | Knŋyan | |
| 66 | 76 | 77 | 78 | 77 | 76 | 74 | 76 | Mstbaun |

Table 5 shows that Palngawan is very different from any of the Squliq dialects. The similarity percentages are so low (64%–70%) that Palngawan can be almost considered a different 'language/dialect' from Atayal. Table 6, on the other hand, shows that Palngawan is also very divergent from other C'uli' dialects. The similarity percentages range from 60% to 69%. This seems to indicate Palngawan is a different 'language/dialect' from other C'uli' dialects. Even Mstbaun is more similar to other C'uli' dialects than is Palngawan (74%–78% vs 66%).

The different word lists yield very different results for the similarity percentages, as reflected in the closeness between Palngawan and other Atayalic dialects in Table 4 and the divergence between Palngawan and other Squliq and C'uli' dialects in Tables 5, 6.

Incidentally, Table 6 seems to provide similarity percentages that support Li's (1996: 192–193) findings that Mkgugut, Pyahaw, Ryuhiŋ, Mtlaŋan and Knŋyan are very similar to one another (93%–96%) but are very different from Mnibu? and Mnawyan (66%–71%). Meanwhile, Mnibu? and Mnawyan are very similar to each other (96%). This also seems to support Li's (1981) claim that Squliq dialects are fairly uniform while C'uli' dialects are considered to be the most divergent.

4 We use *Palngawan* rather than *Palŋawan* to retain consistency in spelling in the paper.

However, since similarity percentages have been responsible for so much confusion, as pointed out correctly by Grimes (1995) — appearing to be easy to calculate and understand whereas riddled with difficulties — we can not draw any conclusions based on shared vocabulary counting alone. Instead, it is better to use cognate strength, calculated by the COMPASS program, to obtain the answer to our first research question whether Palngawan is more similar to Atayal or Sediq.

### *4.2.2 COMPASS results*

The 'COMPASS' function was used to produce an item pairs list calculating cognate strengths and tables of number of words, within given ranges of strength, between three pairs of languages/dialects: Mstbaun-Palngawan, Palngawan-Inago, and Mstbaun-Inago. The tables for the first pair are generated from both word lists (326 words, 190 words) while those for the last two pairs are from the first word list (326 words). The results are presented in Tables 7–11. True cognates are first chosen from the word lists based on the highest average segment strength, 1.00, in the COMPASS tabulation, followed by those with strengths between 0.95 and 1.00, between 0.90 and 0.95, and between 0.85 and 0.90.

#### *4.2.2.1 True cognates (strength = 1)*

**Table 7:** True cognates (strength = 1) among Mstbaun, Palngawan, and Inago based on the word list of 326 lexical items

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 166 | *pila?* | *pila?* | *pila?* | money |
| 325 | *?isu?* | *?isu?* | *?isu?* | you (sg.) |
| 122 | *hiya?* | *hiya?* | *hiya?* | he |
| 322 | *msuyak* | *masurak* | *msurak* | yawn |
| 275 | *ŋuŋu?* | *ŋuŋu?* | *ŋuŋu?* | tail |
| 313 | *?ima?* | *?ima?* | *?ima?* | who |
| 312 | *?inu?* | *?inu?* | *?inu?* | where |
| 309 | *mhuyiq* | *mahuri?* | *mhuriq* | wet |
| 232 | *mpitu?* | *mapitu?* | *mpitu?* | seven |
| 305 | *?ita?* | *?ita?* | *ita?* | we (incl.) |
| 104 | *hi?* | *hi?* | *hii?* | flesh, meat |
| 186 | *?ini?* | *?ini?* | *?ini?* | not |
| 62 | *?ina?* | *?ina?* | *?ina?* | daughter-in-law |
| 226 | *bnaqiy* | *buna?iy* | *bnaqiy* | sand |
| 260 | *taŋuw* | *taŋuw* | *taŋuw* | sprout |
| 298 | *pipi?* | *pipi?* | *pipi?* | vulva |
| 295 | *qalaŋ* | *?alaŋ* | *?alaŋ* | village |
| 220 | *malah* | *malah* | *malah* | to warm |
| 53 | *lukus* | *lukus* | *lukus* | clothes |
| 91 | *mtakuy* | *matakur* | *mtakur* | fall |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 269 | *habuk* | *habuk* | *habuk* | straps |
| 268 | *btunux* | *batunux* | *btunux* | stone |
| 250 | *mhnuk* | *mahnuk* | *mhnuk* | soft |
| 266 | *sknux* | *sakanux* | *sknux* | stink |
| 123 | *tunux* | *tunux* | *tunux* | head |
| 115 | *musaʔ* | *musaʔ* | *musaʔ* | go |
| 121 | *lubuw* | *lubuw* | *lubuw* | jew's-harp |
| 281 | *pqayaʔ* | *paʔayaʔ* | *pqayaʔ* | hang down |
| 43 | *bliŋ* | *baliŋ* | *bliŋ* | cave, hole |
| 278 | *tmalaŋ* | *tumalaŋ* | *tmalaŋ* | taste |
| 40 | *mlawaʔ* | *malawaʔ* | *mlawaʔ* | call |
| 192 | *kuluʔ* | *kuluʔ* | *kuluʔ* | pail, box |
| 242 | *syaw* | *syaw* | *siyaw* | side |
| 267 | *lhbun* | *lahabun* | *lhbun* | stomach |
| 21 | *habuk* | *habuk* | *habuk* | belt |
| 66 | *paraʔ* | *paraʔ* | *paraʔ* | deer |
| 36 | *smayuk* | *sumaruk* | *smaruk* | broil |
| 29 | *qasuʔ* | *ʔasuʔ* | *ʔasuʔ* | boat |
| 96 | *qnalaŋ* | *ʔinalaŋ* | *qnalaŋ* | fence |
| 157 | *gituʔ* | *gituʔ* | *gituʔ* | loquat |
| 156 | *bgiyaʔ* | *bagiraʔ* | *bgiyaʔ* | reed of loom |
| 214 | *mbinah* | *mubinah* | *mbrinah* | return |
| 52 | *galiq* | *galiʔ* | *galiq* | cloth |
| 110 | *raŋiʔ* | *raŋiʔ* | *daŋiʔ* | friend |
| 85 | *gbyan* | *gabyan* | *gbiyan* | evening |
| 15 | *ŋurus* | *ŋurus* | *ŋudus* | beard |
| 154 | *prahuŋ* | *parahuŋ* | *pdahuŋ* | lips |
| 162 | – | *habaraw* | *hbaraw* | many (people) |
| 320 | – | *matas* | *matas* | write |
| 160 | – | *lalbuʔ* | *llbuʔ* | low |
| 77 | – | *bicuw* | *bicuw* | earthworm |
| 293 | – | *tarasiʔ* | *tarasiʔ* | umbrella |
| 270 | – | *mŋaŋah* | *mŋaŋah* | stupid |
| 152 | – | *raklic* | *raklic* | leopard |
| 145 | – | *mapikaʔ* | *mpikaʔ* | lame |
| 14 | – | *balukuʔ* | *blukuʔ* | winnowing |
| 273 | – | *lumaŋuy* | *lmaŋuy* | swim |
| 200 | – | *haruŋ* | *haruŋ* | pine tree |
| 16, 207 | – | *mabatunux* | *mbtunux* | beautiful, lovely |
| 9 | – | *batakan* | *btakan* | bamboo |
| 8 | – | *ʔabulic* | *qabulic* | ashes |
| 261 | – | *rapic* | *rapic* | flying squirrel |
| 81 | – | *maspac* | *maspac* | eight |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 323 | *kawas* | – | *kawas* | year |
| 158 | *sumiq* | – | *sumiq* | body louse |
| 314 | *labaŋ* | – | *labaŋ* | wide |
| 75 | *rhyan* | – | *dhran* | earth |
| 234 | *sasaw* | – | *sasaw* | shade |
| 306 | *tminun* | – | *tminun* | weave |
| 67 | *libuʔ* | – | *libuʔ* | den, nest |
| 265 | *mrŋin* | – | *mdŋin* | sticky |
| 302 | *qsyaʔ* | – | *qsiyaʔ* | water |
| 184 | *tmatuk* | – | *tmatuk* | nod head |
| 296 | *qsahuy* | – | *qsahur* | mind, inner heart |
| 11 | *yawaʔ* | – | *rawaʔ* | bamboo basket |
| 217 | *balay* | - | *balay* | right (correct) |
| 94 | *qthuy* | – | *qthur* | fat, rough |
| 54 | *yuluŋ* | – | *ruluŋ* | cloud |
| 289 | *raŋay* | – | *daŋar* | trap |
| 127 | *sulay* | – | *sulay* | anus |
| 284 | *wayay* | – | *waray* | thread |
| 165 | *karaŋ* | – | *kadaŋ* | molar |
| 280 | *qayaʔ* | – | *qayaʔ* | thing |
| 258 | *takuʔ* | – | *takuʔ* | spoon, scoop |
| 113 | *ʔutux* | – | *ʔutux* | ghost |
| 271 | *bagan* | – | *rbagan* | summer |
| 112 | *qluŋ* | – | *qluŋ* | edible fungus |
| 188 | *smuran* | – | *smudan* | old thing |
| 106 | *phpah* | – | *phpah* | flower |
| 185 | *stunux* | – | *stunux* | noisy |
| 204 | *siyaŋ* | – | *siyaŋ* | pork |
| 100 | *tubaʔ* | – | *tubaʔ* | fish-poison |
| 138 | *kuʔ* | – | *kuʔ* | I |
| 98 | *puniq* | – | *puniq* | fire |
| 256 | *tuyuq* | – | *tuyuq* | spittle |
| 173 | *slaq* | – | *slaq* | mud |
| 5 | *smyuk* | – | *smiyuk* | answer |
| 147 | *msuqiʔ* | – | *msuqiʔ* | late |
| 178 | *pugaʔ* | – | *pugaʔ* | navel |
| 245 | *mtyuʔ* | – | *mtruʔ* | six |
| 227 | *kmugus* | – | *kmugus* | scrub |
| 164 | *skaʔ* | – | *skaʔ* | middle |
| 31 | *pyatuʔ* | – | *pratuʔ* | bowl |
| 42 | *ŋiyaw* | – | *ŋiyaw* | cat |
| 196 | *ʔutas* | – | *ʔutas* | penis |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 193 | *supih* | – | *supih* | ladle |
| 175 | *pupuk* | – | *pupuk* | mumps |
| 46 | *laqi?* | – | *laqi?* | child |
| 44 | *bagah* | – | *bagah* | charcoal |
| 124 | *rmaw* | *rumaw* | – | help |
| 82 | *hiku?* | *hiku?* | – | elbow |
| 161 | *bhluk* | *bahiluk* | – | lungs |
| 308 | *mŋilis* | *maŋilis* | – | weep |
| 153 | *mskkiy* | *maskakiy* | – | to lie on one's side |
| 303 | *tgliq* | *tagli?* | – | waterfall |
| 142 | *buq* | *bu?* | – | juice |
| 297 | *mutaq* | *muta?* | – | vomit |
| 135 | *pira?* | *pira?* | – | how many |
| 134 | *mkilux* | *makilux* | – | hot (weather) |
| 36 | *smayuk* | *sumaruk* | – | broil |
| 35 | *maras* | *maras* | – | bring |
| 2 | *bgayaw* | *bagayaw* | – | Alocasia |
| 285 | *lmuhuw* | *lumuhuw* | – | thread a needle |
| 205 | *limuk* | *limuk* | – | pot |
| 119 | *quri?* | *?uri?* | – | hair, gray |
| 78 | *qpuri?* | *?apuri?* | – | earwax |
| 114 | *miq* | *mi?* | – | give |
| 24 | *mŋihuy* | *maŋihur* | – | salty, hot |
| 111 | *mtŋi?* | *matŋi?* | – | full |
| 192 | *kulu?* | *kulu?* | – | pail, box |
| 259 | *smamaw* | *sumamaw* | – | spread a mat |
| 95 | *mŋuŋu?* | *maŋuŋu?* | – | fear |
| 181 | *sinyuw* | *sinyuw* | – | necklace |
| 79 | *maniq* | *mani?* | – | eat |
| 210 | *mgaliq* | *magali?* | – | ragged |
| 218 | *tuqiy* | *tu?iy* | – | road |
| 252 | *yama?* | *yama?* | – | son-in-law |
| 240 | *qsuyan* | *?asuran* | – | elder sibling |
| 159 | *kuhiŋ* | *kuhiŋ* | – | head louse |
| 190 | *tanux* | *tanux* | – | outside |
| 206 | *ŋahi?* | *ŋahi?* | – | sweet potatoes |
| 33 | *bubu?* | *bubu?* | – | breasts |
| 80 | *tlaqiy* | *tula?iy* | – | eel |
| 41 | *rknus* | *rakinus* | – | camphor laurel |

Table 7 shows that the three dialects share 47 true cognates (strength = 1). Palngawan shares 35 true cognates exclusively with Mstbaun, whereas it shares only 16 true cognates exclusively with Inago. We may draw a tentative conclusion from these numbers that Palngawan shares more cognates with Atayal than Sediq; therefore, Palngawan should be classified as an Atayal dialect, as suggested by Li (1985). However, Mstbaun shares 46 cognates exclusively with Inago but only 35 with Palngawan. It would be misleading to draw the conclusion that Mstbaun should be classified as a Sediq dialect rather than an Atayal dialect.

We then examined the list of cognates at lower strengths of correspondences to find out if the number of shared cognates would change as the threshold is lowered. The results are presented in Tables 8–10 for those with strengths (1) between 0.95 and 1.00, (2) between 0.90 and 0.95, and (3) between 0.85 and 0.90, respectively.

**Table 8:** Comparison of cognates (0.95 <= strength < 1.00) among Mstbaun, Palngawan, and Inago based on the word list of 326 lexical items

(* indicates the item also occurs in Table 7 between different dialects)

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 233 | *smaqis* | *cuma?is* | *sma?is* | sew |
| 194 | *tmapaŋ* | *cumapaŋ* | *smapaŋ* | patch |
| 245 | – | *matu?* | *mtru?* | six* |
| 296 | – | *?acahur* | *qsahur* | mind, heart* |
| 239 | – | *gikus* | *gikus* | shuttle |
| 45 | – | *paskani?* | *paskan* | chew |
| 3 | – | *masa?aŋ* | *masaaŋ* | angry |
| 262 | – | *buhuc* | *brihuc* | squirrel |
| 103 | – | *banux* | *brnux* | flat |
| 60 | – | *kumuc* | *kmruc* | kill |
| 133 | – | *rami?* | *dmai?* | horse |
| 126 | – | *babawi?* | *babaraw* | high |
| 302 | – | *?usye?* | *qsiya?* | water* |
| 235 | – | *mic* | *miric* | sheep |
| 142 | – | *bu?* | *biyuq* | juice* |
| 255 | – | *turu?* | *tudu?* | spine |
| 204 | – | *syeŋ* | *siyaŋ* | pork* |
| 155 | *bsyaq* | – | *busiyaq* | long time |
| 136 | *m?uyay* | – | *mu?uray* | hungry |
| 257 | *tmuyoq* | – | *tmuyuq* | spit |
| 68 | *mhoqin* | – | *mhuqin* | die |
| 38 | *lmoŋ* | – | *lmauŋ* | burn |
| 237 | *qhyaŋ* | – | *hiraŋ* | shoulder |
| 317 | *qmisan* | - | *misan* | winter |
| 197 | *qsyu?* | - | *sru?* | pestle |
| 245 | *mtyu?* | *matu?* | – | six* |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 296 | *qsahuy* | *ʔacahur* | – | mind, heart* |
| 145 | *pikaʔ* | *mapikaʔ* | – | lame |
| 54 | *yuluŋ* | *raruluŋ* | – | cloud* |
| 64 | *kaxaʔ* | *makaxaʔ* | – | day after |
| 113 | *ʔutux* | *ʔamutux* | – | ghost* |
| 270 | *ŋaŋah* | *mŋaŋah* | – | stupid* |
| 47 | *gmoyaw* | *gumuraw* | – | choose |
| 271 | *bagan* | *ʔabagan* | – | summer* |
| 56 | *mumuk* | *ʔumumuk* | – | cover |
| 319 | *smabuʔ* | *cumabuʔ* | – | wrap |
| 264 | *hokuʔ* | *hukuʔ* | – | stick |
| 311 | *knon* | *kanun* | – | when |
| 209 | *ŋahoq* | *ŋahuʔ* | – | pus |
| 291 | *mpusan* | *mapusar* | – | twenty |
| 265 | *mrŋin* | *muraŋir* | – | sticky* |
| 202 | *pturiŋ* | *panturiŋ* | – | point at |
| 225 | *mtnaq* | *mintanaʔ* | – | same |
| 77 | *bisuw* | *bicuw* | – | earthworm* |

**Table 9:** Comparison of cognates (0.90 <= strength < 0.95) among Mstbaun, Palngawan, and Inago based on the word list of 326 lexical items

(@ indicates this item also occurs in Table 8 between different dialects,
* indicates this item also occurs in Table 7 between different dialects)

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 88 | *tquciʔ* | *tiʔutiʔ* | *tquciʔ* | break wind |
| 26 | *qalux* | *makalux* | *mqalux* | black |
| 144 | *tmuciŋ* | *tumutiŋ* | *tmuciŋ* | knock |
| 243 | *ramat* | *raramac* | *damac* | side dish |
| 86 | *quciʔ* | *ʔutiʔ* | *quciʔ* | excrement |
| 167 | *ruŋay* | *ruŋiy* | *ruŋay* | monkey |
| 94 | *qthuy* | *katuhur* | *qtəhur* | fat, rough |
| 197 | *qsyuʔ* | *ʔasuʔ* | *sruʔ* | pestle |
| 8 | *qbuliʔ* | *ʔabulic* | *qabulic* | ashes |
| 224 | *cimuʔ* | *timuʔ* | *cimuʔ* | salt |
| 14 | *lukuʔ* | *balukuʔ* | *blukuʔ* | winnowing |
| 240 | – | *ʔasuran* | *qbsuran* | elder sibling* |
| 2 | – | *bagayaw* | *barayaw* | Alocasia* |
| 284 | – | *wariy* | *waray* | thread* |
| 68 | – | *mahuʔir* | *mhuqin* | die@ |
| 38 | – | *lumoŋ* | *lmauŋ* | burn@ |
| 127 | – | *suliy* | *sulay* | anus* |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 280 | – | *ya?aya?* | *qaya?* | thing* |
| 211 | – | *warux* | *qwarux* | rattan |
| 236 | – | *cumbu?* | *smbu?* | shoot |
| 11 | – | *rarawa?* | *rawa?* | bamboo basket* |
| 124 | – | *rumaw* | *dmayaw* | help* |
| 54 | – | *raruluŋ* | *ruluŋ* | cloud*@ |
| 47 | – | *gumuraw* | *gmaw* | choose@ |
| 303 | – | *tagli?* | *tglaq* | waterfall |
| 225 | – | *mintana?* | *mtna?* | same@ |
| 108 | – | *ciŋas* | *sas* | food particle |
| 155 | – | *buse?* | *busiyaq* | long time@ |
| 4 | – | *sm?ay?aya?* | *smqaya?* | annoyed* |
| 290 | – | *kahnuni?* | *qhuni?* | tree |
| 254 | – | *sinburaŋan* | *smbraŋan* | spear |
| 203 | – | *?arinuc* | *mqrinuc* | poor |
| 57 | – | *rarapa?* | *dapa?* | cow |
| 189 | – | *gumawah* | *rmawah* | open |
| 282 | – | *luŋluŋ* | *lmŋluŋ* | think |
| 291 | – | *mapusar* | *mpusal* | twenty@ |
| 113 | – | *?amutux* | *?utux* | ghost@ |
| 56 | – | *?umumuk* | *gmumuk* | cover@ |
| 175 | – | *tapupuk* | *pupuk* | mumps |
| 271 | – | *?abagan* | *rbagan* | summer*@ |
| 240 | *qsuyan* | – | *qbsuran* | elder sibling* |
| 2 | *bgayaw* | – | *barayaw* | Alocasia* |
| 260 | *kmut* | – | *kmruc* | kill@ |
| 261 | *yapit* | – | *rapic* | flying squirrel* |
| 9 | *takan* | – | *btakan* | bamboo* |
| 70 | *ŋaŋah* | – | *mŋaŋah* | stupid*@ |
| 145 | *pika?* | – | *mpika?* | lame@ |
| 276 | *mlahaŋ* | – | *qmalahaŋ* | take care |
| 316 | *tmabus* | – | *tmbus* | winnow |
| 81 | *mspat* | – | *maspac* | eight* |
| 70 | *spi?* | – | *mspi?* | dream |
| 65 | *qanux* | – | *ruqnux* | deer* |
| 59 | *cyaquŋ* | – | *cyaquŋ* | crow |
| 92 | *waqit* | – | *waqic* | fang |
| 230 | *mita?* | – | *qmita?* | see |
| 262 | *bhot* | – | *brihuc* | squirrel@ |
| 235 | *mic* | – | *miric* | sheep@ |
| 229 | *siluŋ* | – | *wusiluŋ* | sea, lake |
| 152 | *kli?* | – | *raklic* | leopard* |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 260 | *kmut* | *kumuc* | – | kill@ |
| 284 | *wayay* | *wariy* | – | thread* |
| 68 | *mhoqin* | *mahu?ir* | – | die@ |
| 38 | *lmoŋ* | *lumoŋ* | – | burn@ |
| 261 | *yapit* | *rapic* | – | flying squirrel* |
| 127 | *sulay* | *suliy* | – | anus* |
| 280 | *qaya?* | *ya?aya?* | – | thing* |
| 9 | *takan* | *batakan* | – | bamboo* |
| 200 | *hayuŋ* | *haruŋ* | – | pine tree* |
| 81 | *mspat* | *maspac* | – | eight* |
| 293 | *cyasi?* | *tarasi?* | – | umbrella* |
| 136 | *m?uyay* | *ma?uriy* | – | hungry@ |
| 168 | *byaciŋ* | *buratiŋ* | – | moon |
| 326 | *simu* | *cimu* | – | you (pl.) |
| 23 | *kmat* | *kumac* | – | bite |
| 87 | *squci?* | *mas?uti?* | – | defecate |
| 198 | *byok* | *barok* | – | pig |
| 37 | *mumun* | *rumumur* | – | bud |
| 304 | *sami* | *cami* | – | we (excl.) |
| 179 | *sobih* | *sobih* | – | near |
| 109 | *payat* | *parac* | – | four |
| 93 | *tohiy* | *tuhiya?* | – | far |
| 222 | *gamin* | *gamir* | – | root |
| 12 | *tokan* | *tokan* | – | man's basket |
| 301 | *qsya?* | *?usye?* | – | water@ |
| 216 | *pagay* | *pagiy* | – | rice plant |
| 75 | *rhyan* | *rahar* | – | earth* |
| 287 | *hmali?* | *hamalic* | – | tongue |
| 58 | *kagaŋ* | *kakagaŋ* | – | crab |
| 16 | *betunux* | *mabatunux* | – | beautiful* |
| 301 | *mahuq* | *mabahu?* | – | wash (clothes) |

**Table 10:** Comparison of cognates (0.85 <= strength < 0.90) among Mstbaun, Palngawan, and Inago based on the word list of 326 lexical items

($ indicates this item also occurs in Table 9 between different dialects, @ indicates this item also occurs in Table 8 between different dialects, * indicates this item also occurs in Table 7 between different dialects)

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 78 | *qpuri?* | *?apuri?* | *qpuji?* | earwax |
| 311 | *knon* | *kanun* | *knuwan* | when |
| 61 | *mnkuŋ* | *miŋkuŋ* | *mkuuŋ* | dark |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 50 | *mkaraw* | *uŋkaraw* | *mkraw* | climb |
| 103 | *bʔnux* | *banux* | *brnux* | flat |
| 215 | *box* | *box* | *buwax* | rice, husked |
| 236 | *mbuʔ* | *cumbuʔ* | *smbuʔ* | shoot |
| 229 | – | *waciluŋ* | *wusiluŋ* | sea, lake$ |
| 5 | – | *cumik* | *smiyuk* | answer* |
| 92 | – | *wawaʔic* | *waqic* | fang$ |
| 237 | – | *haŋaliʔ* | *hiraŋ* | shoulder@ |
| 12 | – | *tokan* | *tokan* | man's basket$ |
| 143 | – | *putiŋ* | *puciŋ* | small knife |
| 159 | – | *kuhiŋ* | *quhiŋ* | head louse |
| 23 | – | *kumac* | *kmyuc* | bite$ |
| 140 | – | *ʔumbuw* | *rmbuw* | immerse |
| 28 | – | *yumuk* | *miyuk* | blow |
| 136 | – | *maʔuriy* | *muʔuray* | hungry$@ |
| 265 | – | *muraŋir* | *mdŋin* | sticky*@ |
| 20 | – | *nabos* | *nbuyas* | belly |
| 98 | – | *hapuniʔ* | *puniq* | fire* |
| 79 | – | *maniʔ* | *mkan* | eat* |
| 177 | – | *ragirir* | *dgrin* | narrow |
| 120 | – | *maʔas* | *mqaras* | happy |
| 46 | – | *ʔuleʔ* | *laqiʔ* | child* |
| 222 | – | *gamir* | *gamil* | root |
| 12 | *tokan* | – | *tokan* | man's basket$ |
| 133 | *rmeʔ* | – | *dmaiʔ* | horse@ |
| 290 | *qhoniq* | – | *qhuniʔ* | tree$ |
| 282 | *mŋluŋ* | – | *lmŋluŋ* | think$ |
| 124 | *rmaw* | – | *dmayaw* | help*$ |
| 132 | *tryuŋ* | – | *tjiyuŋ* | hornet |
| 208 | *mhoniʔ* | – | *muhnuniʔ* | priest-shaman |
| 146 | *keʔ* | – | *kariʔ* | language |
| 225 | *mtnaq* | – | *mtnaʔ* | same@$ |
| 301 | *mahuq* | – | *mahuʔ* | wash (clothes)$ |
| 228 | *sokiʔ* | – | *sokiʔ* | scythe |
| 200 | *hayuŋ* | – | *haruŋ* | pine tree |
| 83 | *qmyuʔ* | – | *mʔruʔ* | epidemic |
| 45 | *pskon* | – | *paskan* | chew* |
| 195 | *matuk* | – | *gmatuk* | peck |
| 56 | *mumuk* | – | *gmumuk* | cover@$ |
| 285 | *lmuhuw* | – | *lmihuw* | thread a needle |
| 90 | *rqes* | – | *daqras* | face |
| 291 | *mpusan* | – | *mpusal* | twenty@$ |

| Number | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| 141 | *kraya?* | – | *daya?* | inland |
| 163 | *raga?* | – | *dara?* | maple tree |
| 119 | *quri?* | – | *quji?* | hair, grey* |
| 319 | *smabu?* | – | *lmabu?* | wrap |
| 142 | *buq* | – | *biyuq* | juice*@ |
| 264 | *hoku?* | – | *hukuc* | stick@ |
| 229 | *siluŋ* | *waciluŋ* | – | sea, lake$ |
| 5 | *smyuk* | *cumik* | – | answer* |
| 92 | *waqit* | *wawa?ic* | – | fang$ |
| 237 | *qhyaŋ* | *haŋali?* | – | shoulder@ |
| 133 | *rme?* | *rami?* | – | horse@ |
| 290 | *qhoniq* | *kahuni?* | – | tree$ |
| 148 | *kira?* | *kīra* | – | a little later |
| 262 | *bhot* | *buhuc* | – | squirrel*@ |
| 155 | *bsyaq* | *buse?* | – | long time@$ |
| 175 | *pupuk* | *tapupuk* | – | mumps*$ |
| 212 | *mteloq* | *matelu?* | – | raw |
| 13 | *kiri?* | *kagiri?* | – | woman's basket |
| 137 | *qmaluk* | *malrak* | – | hunt |
| 263 | *metaq* | *meta?* | – | stab |
| 84 | *mgey* | *magiy* | – | escape |
| 324 | *hera?* | *hira?* | – | yesterday |
| 152 | *kli?* | *raklic* | – | leopard*$ |
| 318 | *kyu?* | *kuya?* | – | worm |
| 238 | *boluŋ* | *baluluŋ* | – | shrimp |
| 279 | *boq royeq* | *bu?na rori?* | – | tears |

As we lower the threshold to include all cognate pairs with strength ranges above 0.85, the total numbers of cognates exclusively shared between Palngawan–Inago (PA–IN), Mstbaun–Palgnawan (MS–PA), and Mstbaun–Inago (MS–IN) change depending on how the strength level is set.

| Strength | PA–IN | MS–IN | MS–PA |
|---|---|---|---|
| 1.00 | 16 | 46 | 35 |
| ≥0.95 | 31 | 54 | 54 |
| ≥0.90 | 60 | 73 | 85 |
| ≥0.85 | 79 | 98 | 105 |

Therefore, we decided to use only true cognates (strength = 1.00) for comparison. Remember that even this decision is not without arbitrariness. The relationship among the three dialects can be represented in Figure 1.

IN _____ MS _____ PA

**Figure 1**: Dialect chain based on lexical similarity data

This result does not lead us to believe that Palngawan should be classified under either Atayal or Sediq, but rather it shows Palngawan is at the periphery of the Atayalic dialect chain.

A further breakdown of the list of cognates according to their strengths of correspondences can help us define what should be considered exclusively shared cognates between a pair of dialects. Some pairs, indicated with *, @ or $ in Table 10, may change their status from exclusively shared cognates between a pair of dialects to cognates among all three dialects when the strength is lowered; for example, #5 'answer', #92 'fangs', #237 'shoulder', #12 'man's basket', #133 'horse', and #290 'tree'. This will help us evaluate Li's (1985) results with a quantified criterion for exclusively shared cognates.

Before we turn to Li's study, we present the partial results ($0.85 \leq$ strength $\leq 1$) of the COMPASS analysis for the word list of 190 lexical items in Table 11 for comparison.

**Table 11:** Comparison of cognates between Mstbaun and Palngawan based on the word list of 190 lexical items

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| **(strength = 1.00)** | | | |
| 57 | *raŋi?* | *raŋi?* | friend |
| 22 | *?irah* | *?irah* | sister-in-law |
| 133 | *miq* | *mi?* | give |
| 17 | *kira?* | *kira?* | later |
| 115 | *buli?* | *buli?* | small knife |
| 108 | *qasu?* | *?asu?* | boat |
| 45 | *squliq* | *ci?uli?* | person |
| **(0.95 ≤ strength < 1.00)** | | | |
| 75 | *rknus* | *rakinus* | camphor laurel |
| 172 | *sasan* | *sasan* | morning |
| **(0.90 ≤ strength < 0.95)** | | | |
| 21 | *bgira?* | *bagiara?* | batten of loom |
| 54 | *wihiŋ* | *wihiŋ* | water leech |
| 179 | *babaw* | *babaw* | above |
| 186 | *lliw* | *liliw* | tip |
| 58 | *bisuw* | *bicuw* | earthworm |
| 99 | *hmali?* | *hamalic* | tongue |
| 123 | *qmuli?* | *?amulic* | mixed cake |
| 43 | *qbuli?* | *?abulic* | ashes |
| 5 | *sbiŋ* | *cacibiŋ* | sweet |
| 177 | *laxi* | *laxi* | don't |
| 18 | *mqeru?* | *ma?iru?* | nine |

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| **(0.85 ≤ strength < 0.90)** | | | |
| 185 | *kuŋ* | *kuriŋ* | I |
| 128 | *lmoŋ* | *lumoŋ* | burn |
| 32 | *hasa* | *haca?* | there |
| 175 | *suxan* | *cuxan* | tomorrow |
| 66 | *tgliq* | *tagli?* | waterfall |
| 6 | *mtalaŋ* | *matatalaŋ* | run |
| 184 | *ska?* | *cacka?* | between |
| 20 | *kiri?* | *kagiri?* | basket |
| 83 | *taŋuw* | *taŋuw* | bud |
| 16 | *pira?* | *pira?* | how many |
| 174 | *soni?* | *soni?* | today |
| 103 | *lihuy* | *lihul* | forehead |
| 124 | *sbil* | *tasbilian* | lunchbox |
| 120 | *tuqiy* | *tu?iy* | road |
| 42 | *kli?* | *rakalic* | leopard |
| 50 | *ŋli?* | *raŋalic* | fly |
| 73 | *kasi?* | *kamcie?* | sugar |
| 7 | *bagan* | *?abagan* | summer |

### *4.2.3 Comparison with Li (1985)*

Li (1985) claimed Palngawan is an Atayal dialect based on lexical evidence because, among the 800 lexical items, Palngawan shared 214 items exclusively with other Atayal dialects (in his List A), but only 11 items with Sediq (in his List B), and had 50 unique lexical items differing both from other Atayal dialects and from Sediq (in his List C). However, as briefly mentioned in §4, the problems with Li's study are twofold: first, the languages for comparison were not constant but depended on which data were available for comparison; and second, a principle of quantification was not established to group cognates. Our results rectify the two problems.

In Li's (1985) study, Palngawan was compared with Mayrinax as the representative of the Atayal dialects and Tongan as that of the Sediq dialects unless indicated otherwise. As shown in Tables 4, 5 and 6, the similarity percentages within each dialect group (Squliq, C'uli', or Sediq) are not necessarily homogeneous, especially among the C'uli' dialects. It is, therefore, misleading to compare Palngawan with a mixture of Squliq and C'uli' dialects with different similarity percentages, grouped together as 'Other Atayal'.

In Li's List A, where Palngawan is shown to share exclusively with other Atayal dialects but differ from the Sediq dialects, the following lexical items would not be included in the list if different strengths of cognates (< 1.00) were considered.

| Number | MS | PA | IN | Strength | Gloss |
|---|---|---|---|---|---|
| 142 | *buq* | *bu?* | | 1.00 | juice |
| | | *bu?* | *biyuq* | 0.95 | |
| | buq | | *biyuq* | 0.85 | |
| 77 | | *bicuw* | *bicuw*[5] | 1.00 | earthworm |
| | *bisuw*[6] | *bicuw* | | 0.95 | |
| 262 | | buhuc | *brihuc* | 0.95 | squirrel |
| | *bhot* | | *brihuc* | 0.90 | |

When strength is set at 1.00, Palngawan is said to share exclusively with Mstbaun in 'juice' (*bu?* vs *buq*). But if the threshold is lowered to 0.95, Inago can be brought into the cognate set with Palngawan (*bu?* vs *biyuq*). If the strength is again lowered to 0.85, all three dialects can be said to share the same cognate (*buq*, *bu?*, and *biyuq*). By the same token, 'earthworm' and 'squirrel' would not be included in Li's List A.

Similarly, in the same list, pairs that Li cited as exclusively shared cognates also demonstrate different levels of strength as follows:

| Number | MS | PA | IN | Strength | Gloss |
|---|---|---|---|---|---|
| 41 | *rknus* | *rakinus* | | 1.00 | camphor laurel |
| 259 | smamaw[7] | *sumamaw* | | 1.00 | spread a mat |
| 153 | *mskkiy* | *maskakiy* | | 1.00 | to lie on one's side |
| 80 | *tlaqiy* | *tula?iy* | | 1.00 | eel |
| 202 | *pturiŋ* | *panturiŋ* | | 0.95 | point at |
| 198 | *byok*[8] | *barok* | | 0.90 | pig |
| 23 | *kmat* | *kumac* | | 0.90 | bite |
| | | *kumac* | [9]*kmyuc* | 0.85 | |
| 84 | *mgey* | *magiy* | | 0.85 | escape |
| 13 | *kiri?*[10] | *kagiri?* | | 0.85 | woman's basket |

In Li's List C and F where Palngawan is shown to be different from both the other Atayal dialects and Sediq, the following two lexical items would not be included in List C if different strengths of cognates were considered.

5 Li's example is *bicur* in Tongan.
6 Li's example is *bisug* in Mayrinax.
7 Li's example is *sumamag* in Mayrinax.
8 Li's examples are *bauwak*/*?ibubuh* in Mayrinax.
9 Li's example is *qmiyuc* in Tongan.
10 Li's example is *kagisi?* in Mayrinax.

| Number | MS | PA | IN | Strength | Gloss |
|---|---|---|---|---|---|
| 168 | *byaciŋ*[11] | *buratiŋ*[12] | | 0.90 | moon |
| 209 | *ŋahoq* | *ŋahu?*[13] | | 0.95 | pus |

Palngawan shares exclusively with Mstbaun in 'moon' and 'pus' if the strengths are set at 0.90 and 0.95 respectively.

Similarly, the following four lexical items would not be included in List F if different strengths of congates were considered.

| Number | MS | PA | IN | Strength | Gloss |
|---|---|---|---|---|---|
| 237 | *qhyaŋ* | *haŋali?* | *hiraŋ* | 0.85 | shoulder |
| 236 | | *cumbu?* | *smbu?*[14] | 0.90 | shoot |
| 20 | | *nabos* | *nbuyas* | 0.85 | belly |
| 137 | *qmaluk*[15] | *malrak*[16] | | 0.85 | hunt |

'Shoulder' is a cognate among all three dialects at 0.85. Palngawan shares exclusively with Inago in 'shoot' and 'belly' at the level of 0.90 and 0.85 respectively, whereas it shares exclusively with Mstbaun at 0.85 in 'hunt'.

Similarly, based on the results of the COMPASS analysis of the word list of 190 lexical items, partially presented in Table 11, some examples in Li's list A where Palngawan shares exclusively with other Atayal dialects can be ranked according to their strengths as follows.

| Number | MS | PA | Strength | Gloss |
|---|---|---|---|---|
| 45 | *squliq* | *ci?uli?* | 1.00 | person |
| 115 | *buli?* | *buli?* | 1.00 | small knife |
| 5 | *sbiŋ* | *cacibiŋ?* | 0.90 | sweet |
| 175 | *suxan* | *cuxan* | 0.89 | tomorrow |
| 130 | *mtama?* | *tatama?* | 0.77 | sit |
| 104 | *szik* | *sarik* | 0.70 | liver |

But in Li's list F, where Palngawan is different from Squliq and Sediq, the following two examples might be excluded from the list if their strengths as cognates were taken into consideration.

11 Li's example is *buatiŋ* in Mayrinax.
12 Li's example is *baluŋ*.
13 Li's example is *gilu?*.
14 Li's example is *cmebu?*.
15 Li's example is *qmalup*, a form of older generation.
16 Li cited *qumaluap* from Mabatu?an, but no data in Palngawan.

| Number | MS | PA | Strength | Gloss |
|---|---|---|---|---|
| 124 | *sbil* | *tasbilian* | 0.86 | lunchbox |
| 37 | *mu?* | *macmbu?* | 0.54 | shoot |

### 4.3 Summary

Both our shared vocabulary counts and COMPASS analysis seem to, at first glance, point to the conclusion that Palngawan shares more true cognates with Mstbaun than with Inago. But the same method also leads us to the conclusion that Mstbaun is more similar to Inago than Palngawan. If we accept Li's conclusion that Palngawan is closer to Atayal than Sediq based on lexical evidence, and that Palngawan should thus be classified as Atayal, we do not see the whole picture. Only after we understand the lexical similarity between Mstbaun and Sediq can we draw a conclusion on the status of Palngawan. Therefore the results of lexical similarity lead us to believe that, although Palngawan is more similar to Mstbaun than Inago, Mstbaun is also more similar to Inago than Palngawan. In other words, they form a dialect chain, with Mstbaun between Palngawan and Inago. Thus, Palngawan cannot be classified either under Atayal or Sediq but rather at the periphery of the Atayalic dialect chain.
Our investigation of lexical similarity also contributes to further understanding of the problems of interpretation of lexical similarity. We suggested two areas for rectification of the problems in Li (1985). First, the language for comparison should be set constant due to the great divergence among the C'uli' dialects. Second, the criterion for cognate sets can be quantified to reflect relative strengths.

## 5 Sound correspondences

The following section addresses the question whether Palngawan shares more phonological features (i.e. sound change) with Atayal or Sediq. First, the reconstructed phonology and the sound changes that resulted in the contemporary speech varieties are presented. Second, each sound change is documented in detail with actual examples, and data that fail to fit the general patterns are discussed. Finally, four types of sound changes are identified, using Agard's (1984) critieria as further applied in Milliken and Milliken's (1996) work. A rule distribution tableau is made, and correlation coefficients among the sound change rules were calculated to show the relationships among the three Atayalic dialects.

### 5.1 Reconstructed Proto Atayalic phonology

The Proto Atayalic phonology was reconstructed by Li (1981:272), following Dahl's (1976) reconstructed Proto Austronesian phonology, as in Figure 2.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *p* | *t* | | *k* | *q* | *ʔ* | | | |
| *b* | *d* | *g'* | *g* | | | *i* | | *u* |
| | | *c* | | | | | *ə* | |
| | | *s* | *x* | | *h* | | *a* | |
| *m* | *n* | | *ŋ* | | | | | |
| | *l* | *r* | | | | *aw, ay, uy* | | |
| *w* | | *y* | | | | | | |

**Figure 2:** Li's reconstructed Proto Atayalic phonology

### 5.2 Sound correspondences from Proto Atayalic to modern reflexes

Table 12 illustrates Li's (1981) reconstructed Proto Atayalic phonology and the reflexes in Mstbaun, Palngawan, and Inago from our data. Three positions are differentiated, namely word-initial (I), word-medial (M), and word-final (F).

**Table 12:** Sound correspondences in Mstbaun, Palngawan, and Inago

| PA | Modern | PA | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|---|
| **p* | *p-p-p* (I) | **pag'ay* | *pagay* | *pagiy* | *payay* | rice plant |
| | *p-p-p* (M) | **ma-pitu?* | *mpitu?* | *mapitu?* | *mpitu?* | seven |
| | *p-k-k* (F) | **miyup* | *myup* | *yumuk* | *miyuk* | blow |
| | *k-k-k* (F) | **qalup* | *qmaluk* | *malrak* | *maduk* | hunt |
| **t* | *t-t-t* (I) | **tunux* | *tunux* | *tunux* | *tunux* | head |
| | *c-t-c* (I) | **tyaquŋ* | *cyaquŋ* | *te?uŋ* | *cyaquŋ* | crow |
| | *t-t-t* (M) | **kita?* | *mita?* | *tahan* | *qmita?* | see |
| | *c-t-c* (M) | **quti?* | *quci?* | *?uti?* | *quci?* | excrement |
| | *t-c-c* (F) | **waqit* | *waqit* | *wawa?ic* | *waqic* | fang |
| **k* | *k-k-k* (I) | **kadaŋ* | *karaŋ* | *kacaŋ* | *kadaŋ* | molar |
| | *k-k-q* (I) | **kuhiŋ/*kucu?* | *kuhiŋ* | *kuhiŋ* | *quhiŋ* | head louse |
| | *q-k-q* (I) | **kitəhur* | *qthuy* | *katuhur* | *qtəhur* | fat |
| | *k-k-k* (M) | **skənux* | *sknux* | *sakanux* | *skənux* | stink |
| **q* | *q-?-q* (I) | **qabulid* | *qbuli?* | *?abulic* | *qabulic* | ashes |
| | *q-ø-q* (I) | **?ulaqi?* | *laqi?* | *?ule?* | *laqi?* | child |
| | *q-ø-ø* (I) | **qhiraŋ* | *qhyaŋ* | *haŋali?* | *hiraŋ* | shoulder |
| | *q-?-?* (M) | **cumaqis* | *smaqis* | *cuma?is* | *sma?is* | sew |
| | *q-q-q* (F) | **cəlaq* | *slaq* | *calaq* | *səlaq* | mud |
| | *q-?-?* (F) | **mabahuq* | *mahuq* | *mabuhu?* | *mahu?* | wash clothes |
| **?* | *?-?-?* (I) | **?ima?* | *?ima?* | *?ima?* | *?ima?* | who |
| | *?-?-?* (M) | **mu?uray* | *m?uyay* | *ma?uriy* | *mu?uray* | hungry |
| | *?-?-?* (F) | **qudi?* | *quri?* | *?uri?* | *quji?* | gray hair |

| PA | Modern | PA | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|---|
| **b* | *ß-b-b* (I) | **batunux* | *ßtunux* | *batunux* | *btunux* | stone |
| | *k-k-k* (F) | **masurab/ *masuwab* | *msuyak* | *masurak* | *msurak* | yawn |
| **d* | *r-r̃-d* (I) | **daqis* | *rqes* | *r̃ayes* | *daqəras* | face |
| | *r-ʔ-d* (I) | **dapal* | *rapan* | *ʔapar* | *dapin* | sole |
| | *r-r-d* (M) | **pada?* | *para?* | *para?* | *pada?* | pygmy deer |
| | *r-r-j* (M) | **qudi? *qudas* | *quri?* | *ʔuri?* | *quji?* | grey hair |
| | *ʔ-c-c* (F) | **qawlid* | *qoli?* | *ʔolic* | *qowlic* | mouse |
| **g* | *ɣ-g-g* (I) | **gamil* | *ɣamin* | *gamir* | *gamin* | root |
| | *ɣ-g-g* (M) | **kumugus* | *kmuɣus* | *kakugus* | *kmugus* | scrub |
| | *ø-ø-r* (M) | **qagum* | *qom* | *ʔoŋ* | *ʔaruŋ* | anteater |
| | *ø-ø-ø* (M) | **kagac* | *kmat* | *kumac* | *kmyuc* | bite |
| | *ɣ-ø-r* (M) | **daga?* | *raɣa?* | *r̃a?* | *dara?* | Maple tree |
| | *y-y-y* (F) | **bunaqig* | *bnaqiy* | *buna?iy* | *bnaqiy* | sand |
| | *w-w-w* (F) | **lubug* | *lubuw* | *lubuw* | *lubuw* | jew's-harp |
| **g*[17] | *ɣ-g-r* (M) | **bagayag* | *bɣayaw* | *bagayaw* | *barayaw* | Alocasia |
| **g'* | *r-r-y* (M) | **pig'a?* | *pira?* | *pira?* | *piya?* | how many |
| | *y-r-y* (M) | **bagig'a?* | *bgiya?* | *bagira?* | *bgiya?* | reed of loom |
| | *r-r-g* (M) | **cuhig'a?* | *hera?* | *hira?* | *siga?* | yesterday |
| | *y-y-y* (F) | **mabarig'* | *baziy* | *miniy* | *mariy* | buy |
| | *y-y-y* (F) | **kəgig'* | *kgiy* | *kuŋkagiy* | *kəriy* | hemp |
| **c* | *s-c-s* (I) | **cəlaq* | *slaq* | *calaq* | *səlaq* | mud |
| | *s-c-s* (M) | **qacahur* | *qsahuy* | *ʔacahur* | *qsahur* | inner heart |
| | *ʔ-ʔ-c* (F) | *ʔhawkuc* | *hoku?* | *huku?* | *hukuc* | stick |
| | *t-c-c* (F) | **kuməguc* | *kmut* | *kumuc* | *kmruc* | kill |
| **s* | *s-s-s* (I) | **siyag* | *syaw* | *syaw* | *siyaw* | side |
| | *ø-ø-s* (I) | **səpat* | *payat* | *parac* | *səpac* | four |
| | *s-s-s* (M) | **mapusal* | *mpusan* | *mapusar* | *mpusan* | twenty |
| | *s-s-s* (F) | **lukus* | *lukus* | *lukus* | *lukus* | clothes |
| **x* | *x-x-x* (M) | **makaxa?* | *kaxa?* | *makaxa?* | *ŋkaxa?* | day after tomorrow |
| | *x-x-x* (F) | **tunux* | *tunux* | *tunux* | *tunux* | head |
| **h* | *h-h-h* (I) | **hii?* | *hi?* | *hi?* | *hii?* | flesh |
| | *ø-h-ø* (I) | **hapuy* | *puniq* | *hapuni?* | *puniq* | fire |
| | *h-h-h* (M) | **mabahuq* | *mahuq* | *mabuhu?* | *mahu?* | wash clothes |
| | *h-h-h* (F) | **malah* | *malah* | *malah* | *malah* | to warm |
| **r*[18] | *ø-ø-r* (M) | **kari?* | *ke?* | *ke?* | *kari?* | language |
| | *z-ø-g* (M) | **ʔiril* | *tzin* | *ʔir* | *ʔigiy* | left |

17 This **g* corresponds to Proto Austronesian **y*.

18 This **r* corresponds to Proto Austronesian **ɣ*.

| PA | Modern | PA | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|---|
| *r | *y-r-r* (I) | **ruluŋ* | *yuluŋ* | *raruluŋ* | *ruluŋ* | cloud |
| | *ø-r-l* (I) | **rima?* | *imagan* | *ramagar* | *lima?* | five |
| | *ø-r-r* (I) | **raŋaw* | *ŋli?* | *raŋalic* | *rəŋəci?* | fly |
| | *y-r-r* (M) | **qəbəsuran* | *qsuyan* | *?asuran* | *qbsuran* | elder sibling |
| | *z-n-r* (M) | **mabarig'* | *baziy* | *miniy* | *mariy* | buy |
| | *y-ø-r* (M) | **matəru?* | *mtyu?* | *matu?* | *mtəru?* | six |
| | *y-r-r* (F) | **kitəhur* | *qthuy* | *katuhur* | *qtəhur* | fat |
| *l[19] | *l-l-l* (I) | **lumuhug* | *lmuhuw* | *lumuhuw* | *lmihuw* | thread a needle |
| *l | *l-l-l* (M) | **pila?* | *pila?* | *pila?* | *pila?* | money |
| | *ø-ø-ø* (M) | **qaliutux* | *?utux* | *?amutux* | *?utux* | ghost |
| | *n-r-n* (F) | **mapusal* | *mpusan* | *mapusar* | *mpusan* | twenty |
| *n | *n-n-n* (M) | **tinun* | *tminun* | *tuminu?* | *tminun* | weave |
| *m | *m-m-m* (I) | **mataq* | *mteloq* | *matelu?* | *mi?iluq* | raw |
| | *m-m-m* (M) | **?ima?* | *?ima?* | *?ima?* | *?ima?* | who |
| | *ŋ-ŋ-ŋ* (F) | **padahum* | *prahuŋ* | *parahuŋ* | *pdahuŋ* | lips |
| *ŋ | *ŋ-ŋ-ŋ* (I) | **ŋuŋu?* | *ŋuŋu?* | *ŋuŋu?* | *ŋuŋu?* | tail |
| | *ŋ-ŋ-ŋ* (F) | **kadaŋ* | *karaŋ* | *kacaŋ* | *kadaŋ* | molar |
| *w | *w-w-w* (I) | **waray* | *wayay* | *wariy* | *waray* | thread |
| | *w-w-w* (M) | **rawa?* | *yawa?* | *rarawa?* | *rawa?* | bamboo basket |
| *y | *y-y-y* (M) | **qəya?* | *qaya?* | *ya?aya?* | *qaya?* | thing |
| *a | *a-a-a* (M) | **?ita?* | *?ita?* | *?ita?* | *?ita?* | we (incl.) |
| | *a-a-a* (M) | **caqis* | *smaqis* | *cuma?is* | *sma?is* | sew |
| | *ø-a-a* (M) | **laŋuy* | *mŋyoq* | *lumaŋuy* | *lmaŋuy* | swim |
| *i | *i-i-i* (M) | **?inu?* | *?inu?* | *?inu?* | *?inu?* | where |
| | *i-i-i* (M) | **daŋi?* | *raŋi?* | *raŋi?* | *daŋi?* | friend |
| *u | *u-u-u* (M) | **kucu?* **kuhiŋ* | *kuhiŋ* | *kuhiŋ* | *quhiŋ* | head louse |
| | *u-u-u* (M) | **?isu?* | *?isu?* | *?isu?* | *?isu?* | you (sg.) |
| *ə | *ə-u-ə* (M) | **bənaqig* | *bənaqiy* | *buna?iy* | *bənaqiy* | sand |
| | *ø-ø-ə* (M) | **rakəlid* | *kli?* | *raklic* | *rakəlic* | leopard |
| *aw | *o-o-o* (M) | **tawkan* | *tokan* | *tokan* | *tokan* | man's basket |
| | *o-u-o* (M) | **manahawqil* | *mhoqin* | *mahu?ir* | *mhoqin* | die |
| | *aw-aw-aw* (F) | **babaw* | *babaw* | *bawi?* | *baraw* | above |
| *ay | *e-e-e* (M) | **maytaq* | *metaq* | *meta?* | *metaq* | stab |
| | *ay-iy-ay* (F) | **pag'ay* | *pagay* | *pagiy* | *payay* | rice plant |
| *ai | *e-e-ai* (M) | **suwai?* | *sswe?* | *suse?* | *swai?* | younger sibling |
| *uy | *u-u-u* (M) | **kahuy* | *qhuniq* | *kahuni?* | *qhuni?* | tree |
| | *y-uy-uy* (F) | **laŋuy* | *mŋyoq* | *lumaŋuy* | *lmaŋuy* | swim |
| *ø[20] | *ø-ø-ø* (M) | **sauk* | *smok* | *suŋkanux* | *pskənux* | smell |

19 This **l* corresponds to Proto Austronesian **c*.

20 This *ø* corresponds to Proto Austronesian **g'*.

## 5.3 Sound change in the Atayalic group

We begin our discussion with the sound changes that are characterised by Li (1980) as genetically shared by most dialects in the entire group. Some sound changes are completed changes (§5.3.1), while others are ongoing changes (§5.3.2). Sound correspondences that occurred three or more times in our COMPASS analysis are included in our discussion.

### *5.3.1 Completed changes*

*5.3.1.1* -t, -d > -c

In Palngawan and Inago, *-t, -d > -c*:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **waqit* | *t-c-c* | *waqit* | *wawaʔic* | *waqic* | fang |
| **qawlid* | *ʔ-c-c* | *qoliʔ* | *ʔolic* | *qowlic* | mouse |

*5.3.1.2* -b > -p

No examples that illustrate the rule *-b > -p* can be found in our data because most of them have further undergone *-p > -k*.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **masurab/ *masuwab* | *k-k-k* | *msuyak* | *masurak* | *msurak* | yawn |

*5.3.1.3* *-g > -w, -y

Final **-g* becoming *-w* or *-y* in all three dialects is illustrated below:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **bunaqig* | *y-y-y* | *bnaqiy* | *bunaʔiy* | *bnaqiy* | sand |
| **lubug* | *w-w-w* | *lubuw* | *lubuw* | *lubuw* | jew's-harp |

*5.3.1.4* c > s

In Mstbaun and Inago, *c > s*:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **cəlaq* | *s-c-s* | *slaq* | *calaq* | *səlaq* | mud |
| **qacahur* | *s-c-s* | *qsahuy* | *ʔacahur* | *qsahur* | inner heart |

### *5.3.2 Ongoing changes*

*5.3.2.1* -l > -n

In both Mstbaun and Inago, *-l > -n*, but in Palngawan, *-l > -r* (retroflex *r*).

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **mapusal* | *n-r-n* | *mpusan* | *mapusar* | *mpusan* | twenty |

There are variations between *-l* and *-n* as illustrated by the following list of words (Table 13) from our preliminary fieldwork in Mstbaun. Age was found to influence the direction of change (Li 1982a). But further investigation is needed to determine whether the variation is due to lexical diffusion or conditioned by other phonological and social factors.

**Table 13:** Variations between *-l* and *-n* in Mstbaun

| Gloss | Informant 1 (Y.P. 65 years old) | Informant 2 (B.T. 34 years old) |
|---|---|---|
| hundred | *kbhol* | *kbhol* |
| lunch box | *sbil* | *sbil* |
| bladder | *bubul* | *bubul* |
| earth | *rhyal* | *rhyan* |
| dog | *hozil* | *hozin* |
| Atayal | *tayal, tayan* | *tayal* |
| thirty | *mtyul, mtyun* | *mtyun* |
| pain | *mxan* | *mxan, mxal* |
| woman | *knerin* | *knerin* |
| die | *mhoqin* | *mhoqin* |
| three | *ciwan* | *ciwan* |

In some cases, *-n* is further changed to *-ŋ*:

| Gloss | Informant 1 (Y.P. 65 years old) | Informant 2 (B.T. 34 years old) |
|---|---|---|
| hair | *bukil* | *bukiŋ* |

In the following examples, word medial *-l-* is deleted in all three dialects :

| PA | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|
| **qaliutux* | *ʔutux* | *ʔamutux* | *ʔutux* | ghost |

*5.3.2.2* -p > -k, -m > -ŋ

The changes from final *-p* to *-k* and from final *-m* to *-ŋ* are illustrated in the following examples:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **qalup* | *k-k-k* | *qmaluk* | *malrak* | *maduk* | hunt |
| **padahum* | *ŋ-ŋ-ŋ* | *prahuŋ* | *parahuŋ* | *pdahuŋ* | lips |

Even though the changes are considered complete in Li's (1982) study, our preliminary investigation of a few words with final *-p* and *-m* in Mstbaun (Tables 14, 15) indicates there is still change in progress. There is another variation between *-k* and *-c* in the process of change that requires further investigation.

**Table 14:** Variations between *-p*, *-k* and *-c* in Mstbaun

| Gloss | Informant 1 (Y.P. 65 years old) | Informant 2 (B.T. 34 years old) |
|---|---|---|
| seed | *qhak* | *qhap* |
| blow | *zimuk* | *myup* |
| yawn | *msuyak* | *msuyak* |
| catch | *kmiyak* | *kmiyak* |
| eaves | *talak* | *talak* |
| sink | *tgiyuk* | *tgiyuk* |
| enter | *miyuk* | *miyuk* |
| hunt | *qmaluk* | *qmaluk* |
| opposite shore | *qciyak* | *qciyak* |
| scissors | *qatak* | *qatak* |
| ginger | *qurik* | *qurik* |
| stab | *hmak* | *hmak* |
| suck | *pshuk* | *pshuc* |
| fold | *qmuyuc* | *qmuyuc* |
| to fish | *pnec* | *pnec* |

**Table 15:** Variations between *-m* and *-ŋ* in Mstbaun

| Gloss | Informant 1 (Y.P. 65 years old) | Informant 2 (B.T. 34 years old) |
|---|---|---|
| ant-eater | *qom* | *qom* |
| pork | *syam* | *syaŋ* |
| needle | *rom* | *roŋ* |
| grope | *hŋhuŋ* | *hmhuŋ* |
| gall | *yuhuŋ* | *yuhuŋ* |
| lips | *prhuŋ* | *prahuŋ* |
| taste | *tmalaŋ* | *tmalaŋ* |
| burn | *lmoŋ* | *lmoŋ* |
| run | *mktliuŋ* | *mktliuŋ* |
| wipe | *smoŋ* | *smoŋ* |
| dark | *mnkuŋ* | *mnkuŋ* |

#### *5.3.2.3 Vowel deletion before stress*

Unstressed vowels are always deleted or reduced in Mstbaun, whereas they are retained in Palngawan. Some unstressed vowels are reduced and some are retained in Inago:

| Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|
| *bgayaw* | *bagayaw* | *barayaw* | Alocasia |
| *qbuli?* | *?abulic* | *qabulic* | ashes |
| *mlawa?* | *malawa?* | *mlawa?* | call |
| *bliŋ* | *baliŋ* | *bəliŋ* | cave |
| *smyuk* | *cumik* | *smiyuk* | answer |
| *kmat* | *kumac* | *kmyuc* | bite |
| *smayuk* | *sumaruk* | *smaruk* | broil |
| *kmut* | *kumuc* | *kməruc* | kill |

## 5.4 Sound changes that differentiate Atayal from Sediq

In this section, we discuss the sound changes that are claimed by Li (1980a, 1985) to be major phonological differences between Atayal and Sediq. Sound correspondences that occurred three or more times in our COMPASS analysis are included in our discussion.

### *5.4.1 [-p, -b, -m] versus final [-k, -ŋ]*

Atayal is claimed to retain word-final labial stops and nasals [p, b, m] while Sediq has changed to velars [k, ŋ]. As discussed in §5.3.2.2, Mstbaun retains some final *-p* and *-m* while Palngawan and Sediq have changed completely to velars. In this case, Palngawan behaves more like Sediq than Atayal.

### *5.4.2 [ß, r, ɣ] versus [b, d, g]*

Sediq is claimed to retain voiced stops [b, d, g] whereas Atayal has the corresponding fricatives and liquids [ß, r, ɣ] in word-initial and medial positions. But after a closer look at the correspondences of the reflexes, we need to further divide the rule into three subrules.

Inago and Palngawan retain a voiced stop [b] whereas Mstbaun has the fricative [β] in word-initial and -medial positions.

Inago retains a voiced stop [d] or becomes palatalised to [j] before [i] whereas Mstbaun and Palngawan have changed to [r] in word-medial positions.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **pada?* | *r-r-d* | *para?* | *para?* | *pada?* | pygmy deer |
| **qudi?/*qudas* | *r-r-j* | *quri?* | *?uri?* | *quji?* | grey hair |

In word-initial position, Palngawan has developed trilled *r* [r̃].

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **daqis* | *r-r̃-d* | *rqes* | *r̃ayes* | *daqəras* | face |

Inago and Palngawan retain the voiced stop [g] whereas Mstbaun has fricative [ɣ] in word-initial and -medial positions.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **gamil* | *ɣ-g-g* | *ɤamin* | *gamir* | *gamin* | root |
| **kumugus* | *ɣ-g-g* | *kmuɣus* | *kakugus* | *kmugus* | scrub |

### *5.4.3 /y, z, ø/ versus /r/*

Sediq is claimed to retain /r/ while Atayal has changed to /y/, /z/ or zero. In this case, Palngawan is more similar to Inago than Mstbaun.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **ruluŋ* | *y-r-r* | *yuluŋ* | *raruluŋ* | *ruluŋ* | cloud |
| **qəbəsuran* | *y-r-r* | *qsuyan* | *ʔasuran* | *qbsuran* | elder sibling |
| **kitəhur* | *y-r-r* | *qthuy* | *katuhur* | *qtəhur* | fat |
| **raŋaw* | *ø-r-r* | *ŋli?* | *raŋalic* | *rəŋəci?* | fly |
| **mabarig'* | *z-n*[21]*-r* | *baziy* | *miniy* | *mariy* | buy |

In some cases, only Palngawan retains [r], while Mstbaun and Inago have changed to [ø] and [l] respectively.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **rima?* | *ø-r-l* | *imagan* | *ramagar* | *lima?* | five |

### *5.4.4 /-t, -ʔ/ versus /-c/*

The Proto Atayalic **-d* has reflexes *-t* or *-ʔ* in Atayal, but *-c* in Sediq. This has been discussed in §5.3.1.1. In this case, Palngawan is more similar to Inago than Mstbaun.

### *5.4.5* -g- *versus* -r-

For the Proto Atayalic **-g-* [ɣ], Atayal has *-g-* as reflexes, while Sediq has *-r-*. In this case, Palngawan is more similar to Mstbaun than Inago.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **bagayag* | *ɣ-g-r* | *bgayaw* | *bagayaw* | *barayaw* | Alocasia |
| **qagum* | *ø-ø-r* | *qom* | *ʔoŋ* | *ʔaruŋ* | anteater |

### *5.4.6* -r-, -s-, -r̃- *versus* -y-, -g-

Proto Atayalic **-g'-* was claimed to have reflexes *-r-*, *-s-*, or trill *r* in Atayal, while Sediq generally has *-y-*, if preceded by /i/ *or -g-* elsewhere. In this case, Palngawan is more similar to Mstbaun than Inago.

---

21 The correspondence *z-n-r* occurs less than three times in our data.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **pig'aʔ* | *r-r-y* | *piraʔ* | *piraʔ* | *piyaʔ* | how many |
| **cuhig'aʔ* | *r-r-g*[22] | *heraʔ* | *hiraʔ* | *sigaʔ* | yesterday |

From the discussion above, Palngawan was found to share some phonological features with Atayal and some other features with Sediq. A quantitative analysis of the sound change rules is deferred until §5.6 to address whether Palngawan is more similar to Atayal or Sediq in terms of sound change.

## 5.5 Sound changes that differentiate C'uli' from Squliq

In this section, phonemic and phonetic differences between C'uli' and Squliq dialects, stated in Li (1980a) are presented as a basis for comparison among the three Atayalic dialects.

### *5.5.1 Phonemicisation of /e/ and /o/*

According to Tables 1–3, /e/ and /o/ are both phonemes in Palngawan and Mstbaun, while /e/ is not a phoneme in Inago.

### *5.5.2 /q/*

Mstbaun retains /q/, wheras Palngawan has merged into /ʔ/.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **qabulid* | *q-ʔ-q* | *qbuliʔ* | *ʔabulic* | *qabulic* | ashes |
| **cumaqis* | *q-ʔ-ʔ* | *smaqis* | *cumaʔis* | *smaʔis* | sew |
| **mabahuq* | *q-ʔ-ʔ* | *mahuq* | *mabuhuʔ* | *mahuʔ* | wash clothes |

In some cases, Palngawan retains /k/, while Mstbaun and Inago have changed to /q/:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **kitəhur* | *q-k-q* | *qthuy* | *katuhur* | *qtəhur* | fat |

### *5.5.3 /z/*

Mstbaun has developed a phoneme /z/, whereas Palngawan and Inago retain /n/ and /r/ respectively, as discussed in §5.4.3.

### *5.5.4 Phonetic features*

#### *5.5.4.1 [β] versus [v]*

The bilabial fricative [β] is observed to occur as labiodental [v] in the speech of younger female speakers of Palngawan.

22 This sound correspondence occurred less than three times in our data.

#### *5.5.4.2 [r] versus [ɹ]*

The liquid /r/ has a contrast between a trilled or flap [r] and retroflex [ɹ] in word-initial position in Palngawan. These are represented as e.g. *ŕa* 'maple tree' (trill/flap) versus *ramagar* 'five' (retroflex).

#### *5.5.4.3 [h] versus [x]*

Mstbaun and Inago have retained the contrast between /h/ and /x/, whereas /h/ is merged with /x/ in Palngawan.

#### *5.5.4.4 Palatalisation*

Palngawan retains [t] before high front vowels while Mstbaun and Inago have undergone palatalisation in word-initial and -medial positions.

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **tyaquŋ* | *c-t-c* | *cyaquŋ* | *teʔuŋ* | *cyaquŋ* | crow |
| **quti?* | *c-t-c* | *quci?* | *ʔuti?* | *quci?* | excrement |

#### *5.5.4.5 Final [r]*

Final *-l* has changed to retroflex *-r* in Palngawan, while Mstabaun and Inago have changed to *-n*, as discussed in §5.3.2.1.

#### *5.5.4.6* -aw, -ay *versus* -ow, -iy

Proto Atayalic **-ay* is raised to *-iy* in Palngawan:

| PA | Correspondence | Mstbaun | Palngawan | Inago | Gloss |
|---|---|---|---|---|---|
| **pag'ay* | *ay-iy-ay* | *pagay* | *pagiy* | *payay* | rice plant |

Palngawan *-aw* is raised to *-ow* in the speech of the younger generation, as shown in the word list of 190 lexical items in Part 2 of Appendix 1.

## 5.6 Classifications of sound change

All the sound changes discussed in §5.3–§5.5 are further classified based on Agard's (1984) critieria, which are further applied in Milliken and Milliken's (1996). A rule distribution tableau is made and listed in Table 16. Type 0 involves only feature change, but the contrastive pattern does not change. In Type 1, one set of dialects loses a contrast, others do not. Thus the overall system of the language still has the contrast. No significant loss of intelligibility is expected from either Type 0 or Type 1. In Type 2, all sets of dialects lose a contrast, which is still unlikely to impede intelligibility. In Type 3, all sets of dialects lose an

earlier contrast. The structural consequences of this loss are different in one set of dialects than in another. At this point, a single underlying representation is no longer possible, thus this is the kind of change that can impede intelligibility for structural reasons. In Type 4, one set of dialects loses a contrast by one route, while another set loses the same contrast by a different route. Since the languages are split apart structurally, such a change normally impedes intelligibility between the two sets of dialects.

**Table 16:** A rule distribution tableau classified according to Agard's criteria

| Rule | Type 0 | Type 1 | Type 2 | Type 3 | Type 4 |
|---|---|---|---|---|---|
| **-g'- > -r-* | | | | x | |
| **-g'- > -y-/ __ i* | | | | x | |
| **-g'- > -g-/* elsewhere | | | | x | |
| phoneme /e/ | | | | x | |
| phoneme /o/ | | | | x | |
| *-b > -p* | | | x | | |
| **-g > -w* | | | x | | |
| **-g > -y* | | | x | | |
| *-l > -n* | | x | | | |
| *-l > -r* | | x | | | |
| *t > c* palatalisation | | x | | | |
| *-t > -c* | | x | | | |
| *-d > -c* | | x | | | |
| *-p > -k* | | x | | | |
| *-m > -ŋ* | | x | | | |
| *c > s* | | x | | | |
| *V > ø /__V́* | | x | | | |
| *d > r* | | x | | | |
| *r > y* | | x | | | |
| *q > ʔ* | | x | | | |
| *h > x* | | x | | | |
| *-ay > iy* | | x | | | |
| *b > β* | x | | | | |
| *r > r̃* | x | | | | |
| *d > j* | x | | | | |
| *g > ɣ* | x | | | | |
| *r > z* | x | | | | |
| *β > v* | x | | | | |
| *-aw > ow* | x | | | | |

Following Milliken's (1988) procedures, we quantified the results of major sound changes established on the basis of word lists to measure the extent of sound change in Atayal (Table 17). Only the first five sound changes belong to Type 3 change where intelligibility can be impeded for structural reasons. For the rest of the sound changes from Type 2 to Type 0, no significant loss of intelligibility is expected.

**Table 17:** Measuring the extent of sound change in Atayal

| Number | Change | Mstbaun | Palngawan | Inago | Type |
|---|---|---|---|---|---|
| 1 | **-g'- > -r-* | 1 | 1 | 0 | 3 |
| 2 | **-g'- > -y-/ __ i* | 0 | 0 | 1 | 3 |
| 3 | **-g'- > -g-/* elsewhere | 0 | 0 | 1 | 3 |
| 4 | phoneme /e/ | 1 | 1 | 0 | 3 |
| 5 | phoneme /o/ | 1 | 1 | 1 | 3 |
| 6 | *-b > -p* | 1 | 0 | 0 | 2 |
| 7 | **-g > -w* | 1 | 1 | 1 | 2 |
| 8 | **-g > -y* | 1 | 1 | 1 | 2 |
| 9 | *-l > -n* | 1 | 0 | 1 | 1 |
| 10 | *-l > -r* | 0 | 1 | 0 | 1 |
| 11 | *t > c* palatalisation | 1 | 0 | 1 | 1 |
| 12 | *-t > -c* | 0 | 1 | 1 | 1 |
| 13 | *-d > -c* | 0 | 1 | 1 | 1 |
| 14 | *-p > -k* | 1 | 1 | 1 | 1 |
| 15 | *-m > -ŋ* | 1 | 1 | 1 | 1 |
| 16 | *c > s* | 1 | 0 | 1 | 1 |
| 17 | *V > ø /__V́* | 1 | 0 | 1 | 1 |
| 18 | *d > r* | 1 | 1 | 0 | 1 |
| 19 | *r > y* | 1 | 0 | 0 | 1 |
| 20 | *q > ʔ* | 0 | 1 | 0 | 1 |
| 21 | *h > x* | 0 | 1 | 0 | 1 |
| 22 | *-ay > iy* | 0 | 1 | 0 | 1 |
| 23 | *b > β* | 1 | 0 | 0 | 0 |
| 24 | *r > r̃* | 0 | 1 | 0 | 0 |
| 25 | *d > j* | 0 | 0 | 1 | 0 |
| 26 | *g > ɣ* | 1 | 0 | 0 | 0 |
| 27 | *r > z* | 1 | 0 | 0 | 0 |
| 28 | *β > v* | 0 | 1 | 0 | 0 |
| 29 | *-aw > ow* | 0 | 1 | 0 | 0 |

Key: 0 = absent, 1 = present; Type = Sound changes based on Agard's criteria (Type 3 to 0)

The correlation coefficient values were calculated for the twenty nine changes from Table 16 and these are presented in Figure 3.

| | Mstbaun | Palngawan | Inago |
|---|---|---|---|
| Mstbaun | 1.00 | | |
| Palngawan | -0.28 | 1.00 | |
| Inago | 0.11 | -0.17 | 1.00 |

**Figure 3:** Dialect similarity matrix based on sound changes

According to this analysis, Mstbaun shares more sound changes with Inago than with Palngawan; therefore, Mstbaun is more similar to Inago than Palngawan. But if we separate Type 3 changes from the rest of the changes and calculate the correlation coefficient based on the five Type 3 changes, Mstbaun and Palngawan are found to share the same sound changes and have a correlation coefficient value of 1.00, as presented in Figure 4. The correlation coefficient values for Type 0 to Type 2 changes are presented in Figure 5. This tells us that Palngawan resembles Mstbaun in Type 3 changes but Mstbaun resembles Inago in the rest of the sound changes.

| | Mstbaun | Palngawan | Inago |
|---|---|---|---|
| Mstbaun | 1.00 | | |
| Palngawan | 1.00 | 1.00 | |
| Inago | -0.67 | -0.67 | 1.00 |

**Figure 4**: Dialect similarity matrix based on Type 3 sound changes

| | Mstbaun | Palngawan | Inago |
|---|---|---|---|
| Mstbaun | 1.00 | | |
| Palngawan | -0.54 | 1.00 | |
| Inago | 0.27 | -0.07 | 1.00 |

**Figure 5:** Dialect similarity matrix based on Type 0 to Type 2 sound changes

## 6 Summary of the findings

### 6.1 Is Palngawan more similar to Atayal or Sediq in terms of lexical evidence?

Contrary to Li's findings that lexical evidence was more useful than phonology for subgrouping Atayal and Sediq, our study shows that lexical counts are riddled with difficulties and confusion. It was also misleading to conclude that Palngawan was an Atayal dialect based on lexical evidence when Palngawan was compared with a mixture of Squliq and C'uli' dialects with different similarity percentages, grouped together as 'other Atayal'. Although the results of shared vocabulary counting and COMPASS analysis indicate that Palngawan shares more cognates with Mstbaun than with Inago, this is only half of the picture. The real story is that Mstbaun, Palngawan and Inago form a dialect chain, with Mstbaun between Palngawan and Inago. In other words, although Palngawan is more similar to Mstbaun than Inago, Mstbaun is also more similar to Inago than Palngawan. Thus Palngawan cannot be simply classified as a C'uli' but rather is another Atayalic dialect at the periphery of the dialect chain.

### 6.2 Does Palngawan share more phonological features (i.e. sound change) with Atayal or Sediq?

There is no evidence from our study to support Li's claim that Palngawan shares more phonological features with Sediq than Atayal. On the contrary, Mstbaun shares more sound changes with Sediq than does Palngawan, consistent with our lexical evidence too. Mstbaun was found to share the majority of sound changes that do not impede intelligibility with Inago while Palngawan was found to share all the Type 3 changes with Mstbaun. In other words, in terms of quantity, Palngawan shares no more phonological features with Atayal than with Sediq, but in terms of quality, Palngawan shares with Atayal the type of sound change that impedes intelligibility with other dialects.

## 7 Dialect intelligibility

In this section, we discuss the results of the dialect intelligibility test. Since intelligibility cannot be predicted from lexical similarity counts (Grimes 1985), a recorded text test (RTT) was conducted to determine dialect intelligibility of the three dialects.

### 7.1 Recorded text test (RTT)

The procedures for RTT follow Casad (1974) and Blair (1990:73–85). A pilot test was conducted with a panel of ten people to finalise the ten questions in each community. The average score for the ten questions was above 90%. Ten adults, both males and females, with an age range from 29 to 82 were chosen as subjects through a network of friends for the final testing. Each listened to a taped introduction, a hometown test (i.e. the subject's own dialect), and two other dialect tapes. Ideally, only testees who scored 100% in the hometown test would be allowed to continue with the rest of the tests. But if one or two testees in a reference site could not reach 100% due to unfamiliarity with the testing precedure, we still allowed them to participate in the study as long as they demonstrated comprehension of the hometown story by retelling it, and the average hometown test scores for the whole group were above 90%. The order for the two other tapes was not randomised, but fixed as follows: (1) MS hometown–PA–IN, (2) PA hometown–IN–MS, (3) IN hometown–PA–MS. A testing session for each testee lasted for approximately 40 to 50 minutes. All answers to the questions were tape-recorded and translated into Chinese for later scoring by the author and two trained graduate assistants. A correct answer was marked with '1', an incorrect with '0', and a half correct with '.5'. The means and standard deviations of the tested scores were used to distinguish inherent intelligibility from acquired intelligibility.

### 7.2 The intelligibility findings

Ten adult testees were recruited from each reference site through a network of friends to ensure an equal distribution of gender and age. Inago was slightly overrepresented by middle-aged testees because male senior citizens were not readily available due to early mortality. Each testee listened to three tapes, including his/her 'hometown' test tape and two other dialect tapes. The three autobiographical stories for each reference sites are transcribed, translated and listed in Appendix 2.

The mean scores between each pair of dialects based on the intelligibility data were calculated and rounded to the nearest integer to indicate their levels of intelligibility, as presented in Figure 6.

| | | *Tapes* | | |
|---|---|---|---|---|
| | | Mstbaun | Palngawan | Inago |
| *Subjects* | Mstbaun | 100 | 16 | 39 |
| | Palngawan | 23 | 98 | 64 |
| | Inago | 32 | 62 | 91 |

**Figure 6:** Summary matrix of the intelligibility data

The average hometown test scores for the three villages are all above 90% (MS = 100, PA = 98, IN = 91). MS subjects scored better in listening to the Inago tape than to the Palngawan tape (MS–IN = 39 vs MS–PA =16). Palngawan subjects also scored higher on the Inago test than on the Mstbaun test (PA–IN = 64 vs PA–MS = 23). Inago subjects scored better on the Palngawan test than on the Mstbaun test (IN–PA = 62 vs IN–MS = 32). The overall intelligibility between Palngawan and Mstbaun is lower than that between Palngawan and Inago, even though intelligibility is not necessarily mutual (MS–PA = 16, PA–MS =23; PA–IN = 64, IN–PA = 62). The higher intelligibility between PA and IN is not unexpected due to their close proximity to each other. IN children walk through PA village to attend the same school. Dialectal contacts and learning are frequent. An IN resident claimed PA speakers were once heard to speak IN dialect when they were drunk.

While PA and IN residents reported they had more contact with each other than with MS, they claimed they should know a little bit of the MS dialect. But only a few MS residents reported they had any contact with IN speakers and almost no one had any contact with PA speakers. PA dialect was thought to be very divergent from any other Atayalic dialect, even by the PA speakers themselves.

We then further calculated the standard deviations for each mean to determine whether the intelligibility scores reflect inherent intelligibility or acquired intelligibility. The former is the degree of understanding a speaker has of a similar variety from the same linguistic stock while the latter is through exposure to it (Blair 1990:24). Blair (1990:25) gives a rule of thumb that if the standard deviation of the intelligibility score is low (less than 10–12%), then the score is probably an indication of inherent intelligibility. If the standard deviation is high (greater than12–15%), then what is being measured is at least partly acquired intelligibility.

The matrix of average intelligibility scores is presented in Figure 7. The standard deviations between each pair of dialects are all high, indicating that a more thorough bilingualism study with reference to the dialects concerned is needed to separate the effects of inherent and acquired intelligibility.

| | | *Tapes* | | | | | |
|---|---|---|---|---|---|---|---|
| | | Mstbaun (N=10) | | Palngawan (N=10) | | Inago (N=10) | |
| | | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| *Subjects* | Mstbaun | 100 | 0.00 | 16.00 | 13.50 | 39.00 | 16.63 |
| | Palngawan | 22.5 | 15.50 | 97.50 | 4.25 | 64.00 | 15.78 |
| | Inago | 32.5 | 17.65 | 61.50 | 19.30 | 91.00 | 9.94 |

**Figure 7:** Matrix of average intelligibility scores

Comparing the average scores with the standard deviations, we found most IN and PA people understand each other's story on the tape well, but some have difficulty. This is what Blair (1990:25) called Situation 1, with both high average score and high standard deviation. For the other two pairs (MS–PA, MS–IN), on the other hand, many people could not understand the story, but a few were able to answer correctly. This is Blair's Situation 3, where the average score is low but the standard deviation is high.

PA has frequent communication with IN but no contact with MS. Some MS speakers claimed to have contact with Tongan dialect (a Sediq dialect closely related to IN), but no contact with PA.

The results of intelligibility testing all show less than 60% intelligibility. According to Blair's (1990:23) criteria, MS, PA, and IN should be referred to either as dissimilar dialects or different languages depending on the conventions governing the use of the terms 'dialect' and 'language' in the area being surveyed. Therefore, Mstbaun, Palngawan and Inago should be considered three dissimilar dialects.

In terms of percentages of intelligibility, Palngawan is more similar to Inago than Mstbaun due to frequent contacts. Mstbaun has higher intelligibility with Inago than with Palngawan. In other words, Palngawan is considered peripheral in the Atayalic intelligibility networks. This finding conforms to the general attitudes among the three speech communities.

## Appendix 1: Word Lists

### Part 1: Word lists with 326 lexical items

Informants' Background:

| Dialect | Informant's Name | Sex | Age |
|---|---|---|---|
| Mstbaun | Batu Temu (Kao, Tsing-hsian) | M | 34 |
| Palngawan | Temi Temu | F | 50 |
| Inago | Tusung Pengan, | M | 47 |
| | Walis Tadaw | M | 55 |

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 1 | *babaw* | *bawi?* | *baraw* | above | #179 |
| 2 | *bgayaw* | *bagayaw* | *barayaw* | Alocasia | #12 |
| 3 | *myuhung* | *masa?aŋ* | *masaaŋ* | angry | |
| 4 | *smqaya?* | *mahubu?* | *smqaya?* | annoyed | |
| 5 | *smyuk* | *cumik* | *smiyuk* | answer | #127 |
| 6 | *kuluŋ* | *máma?*[23] | *qtahi?* | ant | |
| 7 | *qom* | *?oŋ* | *?aruŋ* | anteater | #56 |
| 8 | *qbuli?* | *?abulic* | *qabulic* | ashes | #43 |

23 *mamá?* 'clean, flat land after cultivation'.

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 9 | *takan* | *batakan* | *btakan* | bamboo | |
| 10 | *buqoh* | *ga?iluh* | *blbun* | banana | #15 |
| 11 | *yawa?* | *rarawa?* | *rawa?* | bamboo basket | |
| 12 | *tokan* | *tokan* | *tokan* | man's basket | |
| 13 | *kiri?* | *kagiri?* | *buluŋuy* | woman's basket | #20 |
| 14 | *luku?* | *baluku?* | *bluku?* | winnowing basket | |
| 15 | *ŋurus* | *ŋurus*[24] | *ŋudus* | beard | |
| 16 | *betunux* | *mabatunux* | *mbtunux* | beautiful | |
| 17 | *sakaw* | *pa?* | *halakaw* | bed | |
| 18 | *qpitay* | *katipar* | — | bedbug | |
| 19 | *sigasuk* | *tuŋa?* | *tgak* | belch | |
| 20 | *ktu?* | *nabos*[25] | *nbuyas* | (upper) belly | #35 |
| 21 | *habuk* | *habuk* | *habuk* | belt | |
| 22 | *pzit* | *kabahni?* | *qtuta?* | bird | |
| 23 | *kmat* | *kumac* | *kmyuc* | bite | |
| 24 | *mŋihuy* | *maŋihur*[26] | *mnihur* | bitter, hot, sour | |
| 25 | *qmtux* | *sa?upaŋ* | *qmupaŋ* | bitter | |
| 26 | *qalux* | *makalux* | *mqalux* | black | #170 |
| 27 | *ramu?* | *ramurux* | *dara?* | blood | #105 |
| 28 | *myup* | *yumuk* | *miyuk* | blow | |
| 29 | *qasu?* | *?asu?* | *?asu?* | boat | #108 |
| 30 | *pŋeloq* | *bahuni?* | *buci* | bow | #106 |
| 31 | *pyatu?* | *ratiŋ* | *pratu?* | bowl | |
| 32 | *luqus* | *lu?iŋ* | *luqi?* | brain, marrow | |
| 33 | *bubu?* | *bubu?* | *?unoh* | breasts | |
| 34 | *?aluk* | *huŋu?* | *hakaw* | bridge | |
| 35 | *maras* | *maras* | *matas* | bring | |
| 36 | *smayuk* | *sumaruk* | *smaruk* | broil | |
| 37 | *mumun* | *rumumur* | *tmumun* | bud | |
| 38 | *lmoŋ* | *lumoŋ* | *lmauŋ* | burn | #128 |
| 39 | *baziy* | *miniy* | *mariy* | buy | |
| 40 | *mlawa?* | *malawa?* | *mlawa?* | call | |
| 41 | *rknus* | *rakinus* | *sakus* | camphor laurel | #75 |
| 42 | *ŋyaw* | *ŋaw* | *ŋiyaw* | cat | |
| 43 | *bliŋ* | *baliŋ* | *bəliŋ* | cave, hole | |
| 44 | *bagah* | *beluh* | *bagah* | charcoal | |
| 45 | *pskon* | *paskani?* | *paskan* | chew | |
| 46 | *laqi?* | *?ule?* | *laqi?* | child | |
| 47 | *gmoyaw* | *gumuraw*[27] | *gmaw* | choose | |
| 48 | *tkata?* | *caciryec* | *kjiyac* | cicada | |
| 49 | *mtasaw* | *matasi?* | *mtasaw* | clean | #11 |

24 'hair'.

25 *labos* 'belly' in the speech of younger generation.

26 *maŋihur* 'salty, hot', *sapiser* 'sour'.

27 *gumuraw* (older generation) varies with *gumurow* (younger generation).

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 50 | *mkaraw* | *uŋkaraw*[28] | *mkaraw* | climb | |
| 51 | *qmlu?* | *?unlu?* | *məduk* | close | |
| 52 | *galiq* | *gali?* | *galiq* | cloth | |
| 53 | *lukus* | *lukus* | *lukus* | clothes | |
| 54 | *yuluŋ* | *raruluŋ* | *ruluŋ* | cloud | |
| 55 | *tltu?* | *gahra?* | *lətu?* | cold (thing) | |
| 56 | *mumuk* | *?umumuk* | *gmumuk* | cover | |
| 57 | *kacing* | *rarapa?* | *dapa?* | cow | |
| 58 | *kagaŋ* | *kakagaŋ* | *kmaraŋ* | crab | |
| 59 | *cyaquŋ* | *te?uŋ*[29] | *cyaquŋ* | crow | #48 |
| 60 | *kmut* | *kumuc* | *kməruc* | kill | |
| 61 | *mnkuŋ* | *miŋkuŋ* | *mkuuŋ* | dark | #160 |
| 62 | *?ina?* | *?ina?* | *?ina?* | daughter-in-law | |
| 63 | *ryax* | *r̃ex* | *jiyan* | day | |
| 64 | *kaxa?* | *makaxa?* | *ŋkaxa?* | day after tomorrow | |
| 65 | *qanux* | *wanux* | *ruqənux* | deer | |
| 66 | *para?* | *para?* | *pada?* | deer, pygmy | |
| 67 | *libu?* | *libuk* | *libu?* | den, nest | |
| 68 | *mhoqin* | *mahu?ir* | *mhuqin* | die | #131 |
| 69 | *kmihuy* | *kumehur* | *kmari?* | dig | |
| 70 | *spi?* | *masper* | *msəpi?* | dream | #9 |
| 71 | *mnbuw* | *ma?abu?* | *mimah* | drink | |
| 72 | *turiŋ* | *masturiŋ* | *tujiq* | drip | |
| 73 | *mbusuk* | *manukan* | *bsukan* | drunk | #189 |
| 74 | *papak* | *caŋe?* | *birac* | ear | |
| 75 | *rhyan* | *rahar* | *dhəran* | earth | #61 |
| 76 | *unuw* | *monuw* | *runuw* | earthquake | |
| 77 | *bisuw* | *bicuw* | *bicuw* | earthworm | #58 |
| 78 | *qpuri?* | *?apuri?* | *qpuji?* | earwax | |
| 79 | *maniq* | *mani?* | *məkan* | eat | |
| 80 | *tlaqiy* | *tula?iy* | — | eel | |
| 81 | *mspat* | *maspac* | *masəpac* | eight | |
| 82 | *hiku?* | *hiku?* | *hiqur* | elbow | |
| 83 | *qmyu?* | *?nu?ric* | *mə?əru?* | epidemic | |
| 84 | *mgey* | *magiy* | *qtuliq* | escape | #25 |
| 85 | *gbyan* | *gabyan* | *gbiyan* | evening | |
| 86 | *quci?* | *?uti?* | *quci?* | excrement | |
| 87 | *squci?* | *mas?uti?* | *qmuci?* | defecate | |
| 88 | *tquci?* | *ti?uti?* | *tquci?* | break wind, fart | |
| 89 | *royeq* | *r̃ori?* | *doriq* | eye | |
| 90 | *rqes* | *r̃ayes* | *daqəras* | face | #89 |

28 *uŋkaraw* (older generation) varies with *uŋkarow* (younger generation).

29 *te?uŋ* (older generation) vaires with *teruŋ* (younger generation).

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 91 | *mtakuy* | *matakur* | *mtakur* | fall | |
| 92 | *waqit* | *wawa?ic* | *waqic* | fang | |
| 93 | *tohiq* | *tuhiya?* | *dhiyaq* | far | |
| 94 | *qthuy* | *katuhur* | *qtəhur* | fat, rough | |
| 95 | *mŋuŋu?* | *maŋuŋu?* | *mi?isu?* | fear | |
| 96 | *qnalaŋ* | *?inalaŋ* | *qnalaŋ* | fence | |
| 97 | *tluliŋ* | *taruli* | *luliŋ* | finger | |
| 98 | *puniq* | *hapuni?* | *puniq* | fire | |
| 99 | *qulih* | *?ucix* | *qsurux* | fish | #49 |
| 100 | *tuba?* | *riluŋ* | *tuba?* | fish-poison | |
| 101 | *psabu?* | *kobu?* | *qowbu?* | fishweir | |
| 102 | *imagan* | *ramagar* | *lima?* | five | |
| 103 | *b?nux* | *banux* | *brənux* | flat | |
| 104 | *hi?* | *hi?* | *hii?* | flesh, meat | |
| 105 | *mqliw* | *mulic* | *qluli?* | flow, adrift | #28 |
| 106 | *phpah* | *rapak* | *phəpah* | flower | |
| 107 | *ŋli?* | *raŋalic* | *rəŋəci?* | fly (insect) | #50 |
| 108 | *pspan* | *ciŋas* | *siŋas* | food particles between teeth | #107 |
| 109 | *payat* | *parac* | *səpac* | four | |
| 110 | *raŋi?* | *raŋi?* | *daŋi?* | friend | #57 |
| 111 | *mtŋi?* | *matŋi?* | *mteŋi?* | full | |
| 112 | *qluŋ* | *kakluŋ* | *qəluŋ* | edible fungus | |
| 113 | *?utux* | *?amutux* | *?utux* | ghost | #68 |
| 114 | *miq* | *mi?* | *mu?at* | give | #133 |
| 115 | *musa?* | *musa?* | *musa?* | go | |
| 116 | *haru?* | *ba?iy* | *biqir* | goitre | |
| 117 | *hnunux* | *sinunux* | *snonux* | hair | |
| 118 | *qpugu?* | *?apuhur* | *sala?* | hair whorl | |
| 119 | *quri?* | *?uri?* | *quji?* | hair, grey | |
| 120 | *mqes* | *ma?as* | *mqaras* | happy | |
| 121 | *lubuw* | *lubuw* | *lubuw* | jew's-harp | |
| 122 | *hiya?* | *hiya?* | *hiya?* | he | |
| 123 | *tunux* | *tunux* | *tunux* | head | |
| 124 | *rmaw* | *rumaw*[30] | *dmayaw* | help | |
| 125 | *kgiy* | *kuŋkagiy* | *kəriy* | hemp plant | #24 |
| 126 | *wawiq* | *babawi?* | *baraw* | high | |
| 127 | *sulay* | *suliy* | *sulay*[31] | anus | |
| 128 | *payah* | *pinah* | *parih* | hoe | |
| 129 | *karuh* | *woc* | *bkaruh* | hoe | |
| 130 | *kmyak* | *r̃umili?* | *dmijin* | hold (in hand) | |
| 131 | *hziŋ* | *hiriŋ* | *walu?* | honeybee | |
| 132 | *tryuŋ* | *r̃yuŋ* | *tjiyuŋ* | hornet (bee) | |

30 *rumaw* (older generation) varies with *rumow* (younger generation).

31 *sulay* 'hip', *bliŋ sulay* 'anus'.

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 133 | *rme?* | *rami?* | *dmai?* | horse | |
| 134 | *mkilux* | *makilux* | *mcilux* | hot (weather or thing) | |
| 135 | *pira?* | *pira?* | *piya?* | how many | #16 |
| 136 | *m?uyay* | *ma?uriy* | *mu?uray* | hungry | #162 |
| 137 | *qmaluk* | *malrak* | *maduk* | hunt | #13 |
| 138 | *ku?* | *cu?* | *ku?* | I | |
| 139 | *saku?* | *kuriŋ* | *yaku?* | I | |
| 140 | *tbuw* | *?umbuw* | *rəmbuw* | immerse in water | |
| 141 | *kraya?* | *r̃etux* | *daya?* | inland, up | |
| 142 | *buq* | *bu?* | *biyuq* | juice | |
| 143 | *buli?* | *putiŋ* | *puciŋ* | knife | #115 |
| 144 | *tmuciŋ* | *tumutiŋ* | *tmuciŋ* | knock | |
| 145 | *pika?* | *mapika?* | *mpika?* | lame | |
| 146 | *ke?* | *ke?* | *kari?* | language, word | |
| 147 | *msuqi?* | *sunarahu?* | *msuqi?* | late | #163 |
| 148 | *kira?* | *kir̃a?* | *kiya?* | a little later | |
| 149 | *lbak* | *?abaw*[32] | *wasaw* | leaf | |
| 150 | *sragiy* | *soruk* | *dagic* | leggings | |
| 151 | *tzin* | *?ir* | *?igiy* | left | |
| 152 | *kli?* | *raklic* | *rakəlic* | leopard | #42 |
| 153 | *mskkiy* | *maskakiy* | *skiiy* | lie on one's side | |
| 154 | *prahuŋ* | *parahuŋ* | *pdahuŋ* | lips | |
| 155 | *bsyaq* | *buse?* | *busiyaq* | long time | |
| 156 | *bgiya?* | *bagira?* | *bgiya?* | reed of loom | |
| 157 | *gitu?* | *gitu?* | *gitu?* | loquat | |
| 158 | *sumiq* | *lumi?* | *sumiq* | body louse | #52 |
| 159 | *kuhiŋ* | *kuhi* | *quhiŋ* | head louse | |
| 160 | — | *lalbu?* | *llәbu?* | low | |
| 161 | *bhluk* | *bahiluk* | *baraq* | lung | |
| 162 | *piyux* | *habaraw* | *hbaraw* | many (people) | |
| 163 | *raga?* | *r̃a?* | *dara?* | maple tree | #84 |
| 164 | *ska?* | *cacka?* | *səka?* | middle | |
| 165 | *karaŋ* | *kacaŋ* | *kadaŋ* | molar | #90 |
| 166 | *pila?* | *pila?* | *pila?* | money | |
| 167 | *yuŋay* | *ruŋiy* | *ruŋay* | monkey | |
| 168 | *byaciŋ* | *buratiŋ* | *?itas* | moon | |
| 169 | *luhuŋ* | *lahyuŋ* | *dahuŋ* | mortar | #14 |
| 170 | *yamux* | *r̃imuli?* | *dmuriq* | moss | |
| 171 | *qoli?* | *?olic* | *qowlic* | mouse | |
| 172 | *nqoq* | *ŋawa?* | *quwaq* | mouth | |
| 173 | *slaq* | *calaq* | *səlaq* | mud | #60 |
| 174 | *tlyu?* | *rkinus* | *kliyuc* | mulberry | |

32 *?abaw* (older generation) varies with *?abow* (younger generation).

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 175 | *pupuk* | *tapupuk* | *pupuk* | mumps | |
| 176 | *qehuŋ* | *ihur* | *liwa* | (wild) mushroom | |
| 177 | *gzil* | *ragirir* | *dgərin* | narrow | #169 |
| 178 | *puga?* | *pupuk* | *puga?* | navel | #93 |
| 179 | *sobih* | *sobih* | *labih* | near | |
| 180 | *gryuŋ* | *wariŋ* | *duyuŋ* | neck, back of | |
| 181 | *sinyuw* | *sinyuw* | *wasin* | string | |
| 182 | *roŋ* | *r̃oŋ* | *qomi?* | needle | #116 |
| 183 | — | *gagi?us* | *giyus* | nit | #59 |
| 184 | *tmatuk* | *?umatuŋ* | *tmatuk* | nod head | |
| 185 | *stunux* | *suntatunux* | *stunux* | noisy | |
| 186 | *?ini?* | *?ini?* | *?ini?* | not | |
| 187 | *?uka?* | *?uŋac* | *?uŋac* | not exist | #176 |
| 188 | *smuran* | *sagiraŋan* | *smudan* | old (thing) | |
| 189 | *gmih* | *gumawah* | *rmawah* | open | |
| 190 | *tanux* | *tanux* | *ŋawuc* | outside | |
| 191 | *ŋuyiq* | *ruk* | *ruw* | owl | |
| 192 | *kulu?* | *kulu?* | *kulu?* | box | |
| 193 | *supih* | *tupih* | *supih* | pan | |
| 194 | *tmapaŋ* | *cumapaŋ* | *smapaŋ* | patch | |
| 195 | *matuk* | *?umatuŋ* | *gmatuk* | peck | |
| 196 | *?utas* | *ranah* | *?utas* | penis | #100 |
| 197 | *qsyu?* | *?asú?* | *səru?* | pestle | #117 |
| 198 | *byok* | *barok* | *babuy* | pig | |
| 199 | *?onray* | *saputu?* | *kalac* | pineapple | |
| 200 | *hayuŋ* | *haruŋ* | *haruŋ* | pine tree | #85 |
| 201 | *sgalu?* | *saminalu?* | *pwalu?* | pitiful | |
| 202 | *pturiŋ* | *panturiŋ* | *tumiyu?* | point at | |
| 203 | *qzinut* | *?arinuc* | *mqrinuc* | poor, lonely | |
| 204 | *syaŋ* | *syeŋ* | *siyaŋ* | pork | |
| 205 | *limuk* | *limuk* | *limuk* | pot | |
| 206 | *ŋahi?* | *ŋahi?* | *buŋa?* | sweet potato | |
| 207 | *betunux* | *mabatunux* | *mbtunux* | pretty, lovely, cute | |
| 208 | *mhoni?* | *murahu?* | *muhuni?* | priest-shaman | |
| 209 | *ŋahoq* | *ŋahu?* | *nalaq* | pus | |
| 210 | *mgaliq* | *magali?* | *mhaliq* | ragged | |
| 211 | *qoyux* | *warux* | *qwarux* | rattan | #70 |
| 212 | *mteloq* | *matelu?* | *mi?iluq* | raw | |
| 213 | *mtalah* | *matanah* | *mbanah* | red | |
| 214 | *mbinah* | *mubinah* | *mbrinah* | return | |
| 215 | *box* | *box* | *buwax* | rice, husked | |
| 216 | *pagay* | *pagiy* | *payay* | rice plant | |
| 217 | *balay* | *cubay* | *balay* | right (correct) | |
| 218 | *tuqiy* | *tu?iy* | *?əlu?* | road | #120 |
| 219 | *kituru?* | *kinkahan* | *tuqiy* | road, animal trail | |
| 220 | *malah* | *malah* | *malah* | to warm, roast | |

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 221 | *ramuw* | *ɾamu?iy* | *damux* | roof | #113 |
| 222 | *gamin* | *gamir* | *gamin* | root | |
| 223 | *ŋbay* | *laŋiy* | *halus* | saliva | |
| 224 | *cimu?* | *timu?* | *cimu?* | salt | |
| 225 | *mtnaq* | *mintana?* | *mtəna?* | same | |
| 226 | *bnaqiy* | *buna?iy* | *bnaqiy* | sand | #62 |
| 227 | *kmugus* | *kakugus* | *kmugus* | scrub, shave | |
| 228 | *soki?* | *kawih* | *soki?* | scythe | |
| 229 | *siluŋ* | *waciluŋ* | *wusiluŋ* | sea, lake | #64 |
| 230 | *mita?* | *tahan* | *qmita?* | see | |
| 231 | *ghaq* | *gagrak* | *gəhak* | seed | |
| 232 | *mpitu?* | *mapitu?* | *mpitu?* | seven | |
| 233 | *smaqis* | *cuma?is* | *sma?is* | sew | |
| 234 | *sasaw* | *sasi?* | *sasaw* | shade | #10 |
| 235 | *mit* | *mic* | *miric* | sheep | |
| 236 | *mbu?* | *cumbu?* | *səmbu?* | shoot | #37 |
| 237 | *qhyaŋ* | *haŋali?* | *hiraŋ* | shoulder | #2 |
| 238 | *boluŋ* | *baluluŋ* | *kboluŋ* | shrimp | |
| 239 | *?ikus* | *gikus* | *gikus* | shuttle | |
| 240 | *qsuyan* | *?asuran* | *qbsuran* | elder sibling | |
| 241 | *sswe?* | *suse?* | *swai?* | younger sibling | |
| 242 | *syaw* | *syaw* | *siyaw* | side | |
| 243 | *ramat* | *raramac* | *damac* | side-dish | #122 |
| 244 | *girgin* | *gigiran* | *gigan* | sifter | |
| 245 | *mtyu?* | *matu?* | *mtəru?* | six | |
| 246 | *pulas* | *gagox* | — | skin disease | |
| 247 | *khway* | *matahayuw* | *mthuway* | slow | |
| 248 | *cipoq* | *yuyuk* | *ciway* | small | #154 |
| 249 | *smok* | *suŋkanux* | *pskənux* | smell | |
| 250 | *mhnuk* | *mahnuk* | *mhənuk* | soft | |
| 251 | *rapan* | *?apar* | *dapin* | sole (of foot) | #88 |
| 252 | *yama?* | *yama?* | *?ama?* | son-in-law | |
| 253 | *mhap* | *?unrak* | *gməhak* | sow | |
| 254 | *qoqoq* | *sinburaŋan* | *smbraŋan* | spear | |
| 255 | *qni?turu?* | *turu?* | *tudu?* | spine | |
| 256 | *tuyuq* | *taruna?* | *tuyuq* | spittle | |
| 257 | *tmuyoq* | *pataruna?* | *tmuyuq* | spit | |
| 258 | *taku?* | *hita?* | *taku?* | spoon, scoop | |
| 259 | *smamaw* | *sumamaw*[33] | *smapaw* | spread a mat | |
| 260 | *taŋuw* | *taŋuw* | *taŋuw* | sprout | |
| 261 | *yapit* | *rapic* | *rapic* | flying squirrel | |
| 262 | *bhot* | *buhuc* | *brihuc* | squirrel | |

33 *sumamaw* (older generation) varies with *sumamow* (younger generation).

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 263 | *metaq* | *meta?* | *metaq* | stab | |
| 264 | *hoku?* | *huku?* | *hukuc* | stick, pole | |
| 265 | *mrŋin* | *muraŋir* | *mdəŋin* | sticky | |
| 266 | *sknux* | *sakanux* | *skənux* | stink | |
| 267 | *lhbun* | *lahabun* | *lhəbun* | stomach | |
| 268 | *btunux* | *batunux*[34] | *btunux* | stone | |
| 269 | *habuk* | *habuk* | *habuk* | straps, belt | |
| 270 | *ŋaŋah* | *mŋaŋah* | *mŋaŋah* | stupid | #156 |
| 271 | *bagan* | *?abagan* | *rbagan* | summer | #7 |
| 272 | *wawi?* | *wagi?* | *hedaw* | sun | |
| 273 | *mŋyoq* | *lumaŋuy* | *lmaŋuy* | swim, bathe | |
| 274 | *msaniq* | *pisani?* | *bsaniq* | taboo | |
| 275 | *ŋuŋu?* | *ŋuŋu?* | *ŋuŋu?* | tail | |
| 276 | *mlahaŋ* | *malahaŋ* | *qmalahaŋ* | take care | |
| 277 | *sehuy* | *cehur* | *sari?* | taro | |
| 278 | *tmalaŋ* | *tumalaŋ* | *tmalaŋ* | taste | |
| 279 | *boq royeq* | *bu? na rori?* | *rusuq* | tears | |
| 280 | *qaya?* | *ya?aya?* | *qaya?* | thing | |
| 281 | *pqaya?* | *pa?aya?* | *pqaya?* | hang down | |
| 282 | *mŋluŋ* | *luŋluŋ* | *lmŋəluŋ* | think | |
| 283 | *qani* | *hani* | *nii* | this | #29 |
| 284 | *wayay* | *wariy* | *waray* | thread | |
| 285 | *lmuhuw* | *lumuhuw* | *lmihuw* | thread a needle | |
| 286 | *ciwan* | *tugar* | *toru?* | three | |
| 287 | *hmali?* | *hamalic* | *həma?* | tongue | #99 |
| 288 | *gnux* | *?apanux* | *gupun* | tooth | #86 |
| 289 | *raŋay* | *pinoŋ* | *daŋar* | trap | |
| 290 | *qhoniq* | *ahuni?* | *qhuni?* | tree | |
| 291 | *mpusan* | *mapusar* | *mpusan* | twenty | |
| 292 | *saziŋ* | *sayiŋ* | *daha?* | two | |
| 293 | *cyasi?* | *tarasi?* | *tarasi?* | umbrella, cap | |
| 294 | *?uyiq* | *?ugir* | *?urac* | vein, sinew | #97 |
| 295 | *qalaŋ* | *?alaŋ* | *?alaŋ* | village | |
| 296 | *qsahuy* | *?acahur* | *qsahur* | inner heart | |
| 297 | *mutaq* | *muta?* | *tbərlih* | vomit | |
| 298 | *pipi?* | *pipi?* | *pipi?* | vulva | |
| 299 | *hzinuk* | *hawinuk* | *hginuk* | waist | |
| 300 | *mnaga?* | *mana?* | *tmaga?* | wait | #150 |
| 301 | *mahuq* | *mabuhu?* | *mahu?* | wash (clothes) | |
| 302 | *qsya?* | *?usye?* | *qsiya?* | water | |
| 303 | *tgliq* | *tagli?* | *tgəlaq* | waterfall | #66 |
| 304 | *sami* | *cami* | *yami* | we (exc.) | |
| 305 | *?ita?* | *?ita?* | *?ita?* | we (incl.) | |
| 306 | *tminun* | *tuminu?* | *tminun* | weave | |

[34] *batunux* (older generation) varies with *urati/* (younger generation).

| Number | Mstbaun | Palngawan | Inago | Gloss | Cross-ref. with 190 lexical items list |
|---|---|---|---|---|---|
| 307 | *mqumah* | *rumohak* | — | to work in the field | |
| 308 | *mŋilis* | *maŋilis* | *lmiŋis* | weep | |
| 309 | *mhuyiq* | *mahuri?* | *mhuriq* | wet | |
| 310 | *nanu?* | *?amur* | *manu?* | what | |
| 311 | *knon* | *kanun* | *knuwan* | when | |
| 312 | *?inu?* | *?inu?* | *?inu?* | where | |
| 313 | *?ima?* | *?ima?* | *?ima?* | who | |
| 314 | *labaŋ* | *rahalaŋ* | *llabaŋ* | wide | #168 |
| 315 | *pali?* | *?alihur* | *tərak* | wings | #44 |
| 316 | *tmabus* | *tumapis* | *tmbus* | winnow | #152 |
| 317 | *qmisan* | *mula?iy* | *misan* | winter | |
| 318 | *kyu?* | *kuya?* | *kui?* | worm | |
| 319 | *smabu?* | *cumabu?* | *lmabu?* | wrap | |
| 320 | *miru?* | *matas* | *matas* | write, tatoo | |
| 321 | *mr?uqu?* | *maguruw* | *mqnuqu?* | wrong | |
| 322 | *msuyak* | *masurak* | *msurak* | yawn | |
| 323 | *kawas* | *iŋkaralan* | *kawas* | year | |
| 324 | *hera?* | *hira?* | *siga?* | yesterday | #19 |
| 325 | *?isu?* | *?isu?* | *?isu?* | you (sg.) | |
| 326 | *simu* | *cimu* | *yamu* | you (pl.) | |

**Part 2: Word lists with 190 lexical items**

Informants' Background:

| Dialect | Informant's Name | Sex | Age |
|---|---|---|---|
| Mstbaun | Batu Temu (Kao, Tsing-hsian) | M | 34 |
| Palngawan | Temu Bakan | M | 75 |
| | Bakan Iwal | F | 81 |
| | Api Rupi? | F | 75 |

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 1 | *ŋta?* | *gilung* | chicken |
| 2 | *qhyaŋ* | *haŋali?* | shoulder |
| 3 | *yabux* | *rinaŋ* | sweat |
| 4 | *mbrus* | *ma?ihur* | to lie |
| 5 | *sbiŋ* | *cacibiŋ* | sweet |
| 6 | *mtalaŋ* | *matatalaŋ* | run |
| 7 | *bagan* | *?abagan* | summer |

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 8 | *qora?* | *kora?* | all |
| 9 | *spi?* | *sipil* | dream |
| 10 | *sasaw* | *sasi?* | shade |
| 11 | *mtasaw* | *matasi?* | clean |
| 12 | *bgayaw* | *bagayaw* | Alocasia |
| 13 | *qmaluk* | *macabu?* | hunt |
| 14 | *luhuŋ* | *lahyuŋ* | mortar |
| 15 | *buqoh* | *ga?iluh* | banana |
| 16 | *pira?* | *pira?* | how many |
| 17 | *kira?* | *kira?* | later |
| 18 | *mqeru?* | *ma?iru?* | nine |
| 19 | *hera?* | *hira* | yesterday |
| 20 | *kiri?* | *kagiri?* | women's basket |
| 21 | *bgira?* | *bagira?* | batten of loom |
| 22 | *?irah* | *?irah* | sister |
| 23 | *mbaziy* | *mababiniy* | trade |
| 24 | *kgiran* | *kuŋkagiy* | peel hemp |
| 25 | *mgey* | *magiy* | escape |
| 26 | *pgeran* | *matalayiŋ* | shun |
| 27 | *sragiy* | *soruk* | covering |
| 28 | *mqliw* | *minturu?* | flow |
| 29 | *qaniy* | *kani?* | this |
| 30 | *haniy* | *nil* | here |
| 31 | *qasa* | *kaca?* | that |
| 32 | *hasa* | *haca?* | there |
| 33 | *qitun* | *?atiy* | corn |
| 34 | *qabay* | *?abalic* | jaw |
| 35 | *ktu?* | *labun* | (upper) belly |
| | *hbuw* | *bunax* | (lower) belly |
| 36 | *mihiy* | *mahiy* | beat |
| 37 | *mu?* | *macmbu?* | shoot |
| 38 | *muya?* | *mamuhi?* | plant |
| 39 | *qmun* | *?untaŋ* | swallow |
| 40 | *hozin* | *huyil* | dog |
| 41 | *qoli?* | *?olic* | rat |
| 42 | *kli?* | *rakalic* | leopard |
| 43 | *qbuli?* | *?abulic* | ashes |
| 44 | *pali?* | *?alihul* | wings |
| 45 | *squliq* | *ci?uli?* | person |
| 46 | *bgax* | *bariŋ* | egg |
| 47 | *kwali?* | *ruk* | hawk |
| 48 | *cyaquŋ* | *te?uŋ* | crow |
| 49 | *qulih* | *?ucix* | fish |
| 50 | *ŋli?* | *raŋalic* | fly |
| 51 | *ka?* | *putuc* | mosquito |
| 52 | *sumiq* | *lumi?* | body louse |

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 53 | *qmici?* | *?amagal* | flea |
| 54 | *wihiŋ* | *wihiŋ* | water leech |
| 55 | *?ubu?* | *yuyux* | nest |
| 56 | *qoŋ* | *?om* | pangolin |
| 57 | *raŋi?* | *raŋi?* | friend |
| 58 | *bisuw* | *bicuw* | earthworm |
| 59 | — | *gi?us* | nit |
| 60 | *slaq* | *macalaq* | mud |
| 61 | *rhyan* | *rahal* | earth |
| 62 | *naqiy* | *buna?iy* | sand |
| 63 | *biŋah* | *haŋituh* | star |
| 64 | *siluŋ* | *waciluŋ* | sea, lake |
| 65 | *sbisuw* | *baluŋ* | thunder |
| 66 | *tgliq* | *tagli?* | waterfall |
| 67 | *qsyak* | *tagacaq* | opposite |
| 68 | *utux* | *?amutux* | ghost |
| 69 | *sapiŋ* | *cacapiŋ* | palm tree |
| 70 | *qoyux* | *warux* | rattan |
| 71 | *bonaw* | *tabiŋ* | peanut |
| 72 | *bilus* | *cabilis* | sugar cane |
| 73 | *kasi?* | *kamcie?* | sugar |
| 74 | *?agiq* | *lami?ul* | miscanthus |
| 75 | *rknus* | *rakinus* | camphor laurel |
| 76 | *wasiq* | *ragutumun* | *Solanum nigrum* |
| 77 | *yahuw* | *ragu?* | *Sonchus oleracus* |
| 78 | *bakih* | *bageluh* | *Laportea pterost* |
| 79 | *sqiy* | *iciga?* | nettle, *Urtica thunberg* |
| 80 | *bukin* | *basikal* | *Alpinia speciosa* |
| 81 | *qabaŋ* | *riluk* | *Rubus taiwan* |
| 82 | *qeruk* | *?uli? gehal* | ginger, pepper |
| | *?ulik* | | (hot pepper) |
| 83 | *taŋuw* | *taŋuw* | bud |
| 84 | *raga?* | *ra?* | maple |
| 85 | *hayuŋ* | *tu?iluŋ* | pine |
| 86 | *gnux* | *?apnux* | tooth |
| 87 | *bukiŋ* | *kumis* | body hair |
| 88 | *rapan* | *?apal* | sole |
| 89 | *rqes* | *rayes* | face |
| 90 | *karaŋ* | *kacaŋ* | molar |
| 91 | *mosiq* | *moci royi?* | eye secretion |
| 92 | *puŋih* | *ilis* | tumor |
| 93 | *puga?* | *pupuk* | navel |
| 94 | *bubul* | *yayubun na babus* | bladder |
| 95 | *yaba bahat* | *temubahak* | heart |
| 96 | *qcyan* | *utin* | hip |

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 97 | *ʔuyiq* | *ʔugil* | vein, sinew |
| 98 | *luqus* | *luʔiŋ* | marrow |
| 99 | *hmaliʔ* | *hamalic* | tongue |
| 100 | *ʔutas* | *ranah* | penis |
| 101 | *lliw ʔutas* | *tatukul* | glans penis |
| 102 | *puluc* | *ruruy* | clitoris |
| 103 | *lihuy* | *lihul* | forehead |
| 104 | *szik* | *sarik* | liver |
| 105 | *ramuʔ* | *ramurux* | blood |
| 106 | *pŋeloq* | *paneluk* | bow |
| 107 | *pspan* | *ciŋas* | food particle |
| 108 | *qasuʔ* | *ʔásuʔ* | boat |
| 109 | *tnuxan* | *patanóan* | pillow |
| 110 | *qbubuʔ* | *tamuku* | hat |
| 111 | *muyaw* | *moroŋ* | house |
| 112 | *tatak* | *rasaliʔ* | hut |
| 113 | *ramuw* | *lamuyi*[35] | roof |
| 114 | *qmayah* | *mumarah raŋuʔ* | dry land |
| 115 | *buliʔ* | *buliʔ* | small knife |
| 116 | *roŋ* | *roŋ* | needle |
| 117 | *qsiyuʔ* | *ʔasúʔ* | pestle |
| 118 | *ʔayaŋ* | *ariŋuʔ* | soup |
| 119 | *lupiʔ* | *siru* | mat |
| 120 | *tuqiy* | *tuʔiy* | road |
| 121 | *kagaw* | *cacobah* | broom |
| 122 | *ramac* | *raramac* | side food |
| 123 | *qmuliʔ* | *ʔamulic* | mixed cake |
| 124 | *sbil* | *tasbilian* | lunch box |
| 125 | *thay* | *tahal* | leftover |
| 126 | *maqus* | *tararú* | ask |
| 127 | *smyuk* | *cumik* | answer |
| 128 | *lmoŋ* | *lumoŋ* | burn |
| 129 | *thekan* | *rakac* | stool |
| 130 | *mtamaʔ* | *tatamaʔ* | sit |
| 131 | *mhoqin* | *mahuʔil* | die |
| 132 | *syunaw* | *yumunaw* | substitute |
| 133 | *miq* | *miʔ* | give |
| 134 | *mhŋaw* | *paŋuw* | rest |
| 135 | *surux* | *macaruw* | stand |
| 136 | *qmataq* | *kunteruʔ* | eat raw |
| 137 | *mnayaŋ* | *tumabul* | clear land |
| 138 | *kigaqaw* | *kunagaʔuw* | new land |
| 139 | *lmaguʔ* | *naumic* | keep plants |
| 140 | *lmahiŋ* | *mumarah* | to thin out, to weed |

35 *ramuyi* varies with *ramu/iy*.

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 141 | *mbun* | *abulun* | bury |
| 142 | *mtux* | *manawaral* | bark |
| 143 | *qmatak* | *matak* | cut |
| 144 | *t?asuy* | *masuhul* | cough |
| 145 | *tgyuk* | *maruk* | sink |
| 146 | *rmhaw* | *luŋhaw* | sharpen |
| 147 | *mxan* | *muxal* | pain |
| 148 | *nbu?* | *muxal* | sick |
| 149 | *smxu?* | *cunxu?* | pound |
| 150 | *mnaga?* | *mana?* | wait |
| 151 | *mhkaŋi?* | *makakiy* | walk |
| 152 | *tmabus* | *tumapis* | winnow |
| 153 | *cyaba?* | *yoba?* | big |
| 154 | *cipoq* | *yuyuk* | small |
| 155 | *mbuloq* | *maritux* | blind |
| 156 | *mŋuciq* | *mŋaŋah* | dumb, stupid |
| 157 | *smyax* | *picyeh* | bright |
| 158 | *hiyaq* | *gara?* | cold |
| 159 | *shyu?* | *magaluyiŋ* | straight |
| 160 | *mnkuŋ* | *miŋkuŋ* | dark |
| 161 | *mkyay* | *maŋu?*[36] | dry |
| 162 | *m?uyay* | *ma?uriy* | hungry |
| 163 | *msuqi?* | *sunrahu?* | late |
| 164 | *giqas* | *ga?arus* | new |
| 165 | *mnkis* | *nakis* | old |
| 166 | *msayux* | *masarux* | shy |
| 167 | *mqŋuqu?* | *maŋurah* | sleepy |
| 168 | *glabaŋ* | *rahalaŋ* | wide |
| 169 | *gzil* | *ragiril* | narrow |
| 170 | *qalux* | *makalux* | black |
| 171 | *hmswa?* | *hunco?* | why |
| 172 | *lhŋan* | *cka baŋi?* | night |
| 173 | *sasan* | *sasan* | morning |
| 174 | *soni?* | *soni?* | today |
| 175 | *suxan* | *cuxan* | tomorrow |
| 176 | *?uka?* | *?uŋac* | not have |
| 177 | *laxi* | *laxi* | don't |
| 178 | *hga?* | *laha* | they |
| 179 | *babaw* | *bawi?* | above |
| 180 | *kraya?* | *yatux* | upland |
| 181 | *glaŋ* | *galeŋ* | lead |
| 182 | *suruw* | *bukuy* | behind |
| 183 | *qsahuy* | *rik* | inside |

36 *maŋu?* varies with *maraŋu?*.

| Number | Mstbaun | Palngawan | Gloss |
|---|---|---|---|
| 184 | *ska?* | *cacka?* | between |
| 185 | *lliw* | *liliw* | tip |
| 186 | *kuŋ* | *kuriŋ* | I |
| 187 | *Kuŋ ga Tayal.* | *Itaral kuriŋ*[37] | I am a native. |
| 188 | *Iyat saku Tayal.* | *Arac u Itaral.* | I am not a native (Atayal). |
| 189 | *mbusuk* | *manukan* | drunk |
| 190 | *Nyux ku mbusuk la.* | *Manukan cu la.* | I am drunk. |

## Appendix 2: Recorded text tests for intelligibility

### Mstbaun Text (narrated by Yabu Pawan, M, 65 years old)

*lalu mu Tayal ga Yabu Pawan, lalu Gipun ga Yamagucu Masaaki, lalu Cyukok ga Yang A-Ji. Kawas mu ga mtyu pgan magan.*

*ru arin ku laqi ga mqzinut qu qnxan maku, ru trang keku imagan kawas ga, si say mlahang ku qsuyan maku mbuloq, lalu nya ga Pihaw.*

*ru mqelang myan mcisan lru, kmayan qsuyan maku Pihaw ru: 'Hata hkangi putung ru hata lmom qnahi muci ru', lon myan putung ru, ini saluw hiya, kuzing qu mutung lru, mtlom lru, wan si itta rgyax qu puneq ru, ulung su splawa qora mrkyas qalang ru, son nha muyut.*

*ru gbyan nasa lga, tpihun mina yaba kesat ga bucyow, Kohara bucyow lalu nya. yaya mu ru qsuyan maku Pihaw ru kuzing, muha myan qzitan qnawan (qzitan qnawan hiya ga hasisyo), ru tpihun mina bucyow ru qusan minya ru bhiyan minya. ru bhiyan ku nya ru, ini nya bhziy yan qsuyan maku Pihaw qya mbuloq royeq nya, yaya mu ru kuzing ga bhyan nya.*

*baqun maku Gipun qaniy ga, maki balay regi nya, anay ta tmubun qora ke ru noy ta nya ini bhziy muci saku ru, mihiy lga stubun maku wi maku son 'Konniciwa' ru; mihiy loziy ga 'Kongbangwa' son maku ru, mihiy loziy ga 'Ohayogoraymas' son mu, qora balay regi Gipun ga wan maku skayan ga, ini alay taling mihiy qu Gipun qasa ru son ta alay nanu kmayan soəobey mihiy la i- key nya ini baqiy mung ke ta muci saku ru mihiy loziy, wi knya lklun ru wi knya bhlan ru, qora balay regi wan maku stubun qora ga, ini balay the, obey mihiy.*

*ru yaya maku ga bhyan nya ga imaw mtbuling, ru kmayan yaya mu lga: 'Yat ta pqyanux lru, phoqin ta la, talagay alay mxan hi yayun ta nya mihiy muci yaya mu ru', ke ska bengi lga, pwahun mina bucyow Gipun qasa lru, muha myan ngasan lru.*

*kmayan yaya maku, 'simuw ssekay aku hru, yasa qu hoqin kun muci ru', bhlan nya wasin qolu maku ru, nga ... nga ku mngilis, 'iyat saku balay phci qolu', son maku yaya mu ru, ulung su key wan muci nanu lru, ini ku nya bhiy qolu lru, moyay ku mngilis.*

*nanu yasa qu son mha trang laqi cipoq ga, key kuzing wan nya sqnutan arin nxan na Gipun qaniy. Ima lux baq yow qaniy, Gipun qaniy ga, 'Ini kita bnkis ru laqi, si nya tmahiy tmahiy mihiy ru uka balay ryosin nya qu Gipun qaniy.'*

*ru nanu yasa qu yaya mu uzi ga ulung su key wan muci nanu lru ini nya pskciy qu qolu maku ru. nyux ku mbzinah msmoniy misuw qaniy ga, qaniy ga gnalu na Utux Kayal.*

37 *Itaral kuriŋ* varies with *Itaral cukuŋ*.

**Translation of Mstbaun story**

My Atayalic name is Yabu Pawan. My Japanese name is Yamagucu Masaaki. My Chinese name is Yang, A-ji. I am sixty-five years old.

During my childhood, my family was very poor. When I was about five years old, I began to take care of my older brother, who was blind. His name was Pihaw.

When we were bored with playing, my older brother, Pihaw, suggested, 'Let's go find matches to burn the mountain!' After we found the matches, because he could not see, I was the one who lit the matches to burn the mountain. The fire spread quickly to the top of the mountain. Fortunately, our call summoned all the young people from the village. They came to put out the fire.

That evening, the Japanese policeman called us in. His name was Mr Kohara. My mother, my older brother Pihaw and I went to the police bureau. He interrogated us and then beat us. He beat me but not my older brother Pihaw because he was blind. Thus my mother and I were beaten.

I knew that Japanese had many polite expressions. I thought if I said all the polite words, he might not beat us. When he beat me, I bowed my head and said, 'Good afternoon!' When he beat me again, I said, 'Good evening!' When he still beat me, I said, ' Good morning!' After I had said all the Japanese greetings, he still did not stop beating me. No matter what I said, he still beat me. I wondered if he did not understand what I had said. Therefore, I repeated all the polite expressions again, but it did not work. He still tied me up and beat me.

My mother was beaten so hard that she flew across the room. My mother said, 'We can't live anymore. We'll die. Our bodies have been so badly beaten'. Not until midnight were we released by the policeman. Then we went home.

My mother said, 'I'll first hang you, and then I'll hang myself'. When she tied a rope around my neck, I cried and shouted, 'Don't choke me!' Fortunately, I did not know what happened, but my neck was not tied. But I was still crying very hard.

During my childhood, I was probably the only one who had been beaten hard by the Japanese. Who knows why? Those Japanese did not have any conscience. They beat anybody, young or old.

Something miraculous happened that my mother did not choke me. I have been able to survive until now. This is God's will.

**Palngawan text (narrated by Temi Temu, F, 50 years old)**

*kuring hiya ong Temi Temu ka lalu mu, lalu na itaral kani. ka ausa mu hiya ga, murnarah cu kararih na kanel. ru karuma ga musa mu marah ong, aska kalama musa matoh cu usix. ru amoka usix mu ga maki tagacak na luling.*

*utux rih ong musa cu matoh. moka luling na usix hani ga, tagic si gawah cu yaba, tagic si hatuw. ini mu nak bayi ka kanon karih huntuw uri ru kanon ka gumawah.*

*rih cikaca ga musa cu matoh cu usix tagacak. musa cu hang ga, yuyuk ka luling hani hang, ho ru tamasu cu sumbali cu puying na usix mu hani la ga, bagi cu rasali mu la ga, yaba ka usix la. lunglung cu bayu tanainu mascarak cuka luling hani mikong. ungac carong tatavingan kasun luling hiya kai. aska minoh linglungan tanak hapo mascarak caga. mulәuli kinang ka huyil. maksisiyaw cu ana mu ahkalangi ka pacarapan kaluling hani. hapo ka usix ni. ascu tamaluh macaruw tagacak hang. pasco tala mikong. utux na cu sababaə ka kinanuhan mu hani. pintaringan nak uli ka sababaə ka utux morong min. sababaə min cu ka tensikiu hani ca. lunglung cu yunani pakalu cu rarihung na ausa hani*

*ong. balbali ta cumun yababawi mikong. musa sinbabaəan tanak ong ana ta musa ana inu ga maki babang anali aska ana inu ong. maki kararih kasun yababawi ulasu itan ca ga.*

*ho ru lunglung cu cumun yababawi la ga. ascu lunglung cu ka yaba mu. lima wal bawi karal ka yaba mu. yaba mu sinbali kinang hiya mini ka kinanuhan mu hani. kani ka tagaga cu cumun ga baliə cu lunglung cu ka yaba mu hang. ho ru kumaral cu cuka kei mu. yaba mu isu cebawi. isuka sumbali kinang. yosunani sumbali kinang ong baliun saku pakaiyanux babaw na rahal hani. micu necu bakalu cu ka rarih ong na ausa hani. balii taluhing moka cinun mu hani ga taluhingi paspa yaba bawi ka kei mu hani. kani ka lamun mu kora kasun kei na kiukai sinbabaan mu nak hiya ga, tensikiu kani ka lungpu cu tianzhu jing, shengmu jing, guangrong jing.*

*tamasu cu cumun yaba bawi la ga, ho mukung puson mu kora ka ugil mu tanabuy cu yaba bawi pinlarang cu carong gumoro ka luling hani. asmu lunglungi ka kei ni yaba mu. gomoro luling maka yaba mu ga malahang cu kasun masapow na luling ma. maraara ka usix la ong mahuyo hari ka insa na usix hani. baliun mu malah ka lukus mu. moka huyil mu muuli kinang hani ga kalalama gomoro ka huyil hani. wal mulic nanu hugal ka huril hani kai. ga, iyaə ta lunglungi kaca mikong. tamasu tanak cumun yaba bawi laga. balbali? tanak kunhapo cuka ausa tanak hani. ananak aspaskura babang laga unga ka ciəuliə unil lumo kinang caga. kani ka puson munak kinhapo na linglungan puson munak linglungan ka yaba mu parow kinang ka yaba mu mikong.*

*kani ka gumoro cu ka luling haca la. gomoro cuka luling haca ong. minutux cu umara paragan mu kake mu 'yaba bawi balii cu gunlasu mikong, yaba bawi balii cu gunlasu mikong.' maha cu minramagal matuə gumoro ascu nak tahi mascarak cu kaluling hani la. mascarak cu la ong ascu paksangi sisyaw babang na luling. mayanux ta ong mikong. mayanux tala ga pakalahang ta cu uli ci morong, mrong ta mikong.*

**Translation of Palngawan text**

My name is Temi Temu. This is my Atayaic name. Concerning my life, I go to work every day as a woman. Sometimes I go to the field, but I have to check the source of the water first. The water is on the other side of the river.

One day I went to check the water. The water flow of the river was sometimes very rapid but sometimes under control. I did not know when the water was blocked and when it was released. One day, I went to check the water across the river. When I started, the water flow was small. When I finished checking the source of the water and was on my way back to the hut, the waterflow became rapid. I did not know how to cross the river. I thought there was nothing for me to hold onto in the river. I could only depend on my strong faith to get across. My dog was following me. I was walking near the edge, trying to find a way to cross the river. The waterflow was very strong. I was staring across the river. I was wondering what to do. Fortunately I had faith. Since I was a child, my family have believed in God. We believe in Catholicism. I thought I should pray devoutedly to Our Heavenly Father whenever I came across any difficulty. I believed Our Heavenly Father was protecting us wherever we go at all times.

While I was praying to Our Heavenly Father, I suddenly thought about my father. My father had already gone to Heaven. My father made me and gave me life. Therefore, when I began to pray I thought about my father. I prayed, 'You are my father up there. You made me. Since you made me, you wanted me to live well on earth. Now I am facing a problem in my life. Please pray well for me. Tell my word to Our Heavenly Father'. I remembered all my church teachings and recited Ave Maria, etc.

After I had prayed to Our Heavenly Father, I decided to use up all my energy. I firmly believed in Our Heavenly Father and I was determined to go across the river. All of a sudden, I remembered what my father said before. He said, 'When crossing the river, I look for the shallow place, where the water splits. If the water splits, it has less force'. Thus I took off my clothes. The dog that was following me went ahead of me to cross the river. It was drifting downward. I stopped thinking about my feeling temporarily. Since I had prayed to Our Heavenly Father, I might as well confront it with faith no matter what happened. No matter how much I looked around, hoping to find somebody, no one could come to help me. I decided to strengthen my heart. My father came to mind. I thought my father would help me.

Now I was ready to cross the river. Each step I took, I shouted, 'Heavenly Father, protect me! Heavenly Father, protect me!' I repeated it for five or six times while I was going across the river. Suddenly I saw I was able to cross the river. After I arrived at the bank of the river, I felt totally motionless. I thought, 'I survived'. After I survived, I wanted to take good care of my children and my family.

**Inago text (narrated by Walis Tadaw, M, 55 years old)**

*rngaw mu sayang o nədaəan sapah mu sipiyaw. mensa hetay sipiyaw ka tama mu o taha pi ka mensuwai. kentatah hetay ka tama mu o wata ini baka mssbu ka hetay tanah tunux. ataw ka mensa hetay ka tama mu o mimah sinaw kəjijiyah. muəutux uhway kia o mtjiyan kəjiyah. uhway kia o bubu mu bhragun nia, uhway kia o yami mensuwai maku pi makələmiqu mataqiy. basukan ka hiya naqeh balay. babaw nya o miyah ka təməsaə ka sinəyesu kiokay ka sinəhiyi ka tama mu ta.*

*babaw miyah sinəhiyi ka tama mu ta o ini imah sinaw ta. kia ka sitatao nətrumuc pi miyah kiokay ta. kana bubu mu mi laqi nya kana to miyah kiokay la, rima ka laqi senaw mi truka laqi karijin. kia ka miyah kiokay kana ta. kia ka nədaəan sapah mu sipiyaw.*

*kia ka seikacu nami sipiyaw ta o. blalay piha o mhuma masu, mhuma bunga mi basaw uli. kia ka seikacu sipiyaw kia miyah ka tluw ta o. babaw nya to o seikacu ini tnataə. sipiyaw ho o ini kan abula ha cimu nanak ha kentatah ka menkan abula ta o mi?ing pila ta mhuma masu mi bunga ha. kentatah ka menkan abula ta o, miəing pila ta. mhuma masu mi bunga ha o, ini ingi pila ha. mhuma payay ta o ming pila ta. kia ta o məkələmiqu ta mosa miəing pila ləmiqu. saw kia ta o səmələy kingsering, tsun ta ha sapah. brigun ta ha cimu mi abula. babaw nya o tbriyoh ka seikacu ta. pinaw ka bunga o ini taha kuhi mkan ta itaw masu oli o ini taha kuhi. kia ka sayang ta o miəing pila muyic balay ta. kia ka sayang ta o balalay mhuma nasi ha pinaw yami hini o llubu pika hini kia ka nasi o ini aoli. naqaih ka nasi oli o wata sakatun ta. babaw nasi to o mhuma saka mu la, mhuma saka mu ucula naka mamei oli. babaw nya o mkala ka seikacu babaw thəgani ta o. ana manu ni ta o pila ta o bhrakun ta ha. kia ka seikacu nami ta o mahuma ambali ta. pinaw ambali o basiyaq o maluka netang nia ha. pinaw ka hici ta o, naqeh ka netang ka nia ta. kia ka naqeh ka ambali ta o, musa miəing pila ta. kia ka laqi ta o musa kampah ngawuc la kəmpah katawa ku məgənteking uli. kia ka o ləngəlung nami ta, kia ka parajing miyah təmusa mhuma ocia ta. kia ka mhuma ocia mi betak sayang. yapi zu ko nen ta kenta kia ka seikacu nami ta o, malu hali la. kia ka marana marana marana. kia ka seikacu nami ta ya mhuma nami ocia ni ta o. ana uli o mangan ina uli. uhway kia uli o seikacu laqi ta ha uli o. wata malu malu pika sayang ta. mahuway namu balay ta.*

### Translation of Inago text

Now I am going to talk about what happened to my family. When my father served in the military, there were only my brother and I. He returned home after the Japanese military was defeated in the war. Because he had served in the military, he often drank and sought trouble. After he got drunk, he either punched or chased my mother around. Because there were only two of us as brothers, we went to hide on the mountain whenever he got drunk.

My father was very mean after he drank. But after the Christian church came to preach the Gospel, my father believed in it. When he became a believer, he did not drink anymore. He attended the church regularly and my mother and all the children, five sons and three daughters, went to church. Our whole family went to church. This is what happened to my family.

Now I am going to talk about life in the past. We originally planted millet and sweet potatoes. After the Mainland Chinese came here [Taiwan Retrocession], our life was different. In the past we did not have oil to eat but only salt. If we wanted to eat oil, we had to look for more money. But we did not make much money by planting millet and sweet potatoes; therefore, we went to the mountain to look for another source of income. Sometimes we brought home a kind of herbal medicine [Jin xian lian in Chinese] to sell in order to buy salt and oil. Later the living standard increased more, and our children did not like to eat either sweet potatoes or millet. Thus we were busy trying different ways to make more money. First, we tried to grow pears. But the altitude in this area was too low for pears. So the pear trees were taken down. Next, we tried to grow corns, lima beans and red beans. But the cost of living kept increasing too quickly for us to catch up. It required more money to keep up to par. So we began to grow red pears. The price for the fruit was high for a while. But when the price went too low, we had to go after more money again. So our children left home for construction work, pouring cement and tying steel. Because we were concerned for our children's hard labor, we began to grow tea. It has been fifteen years since we started to grow tea. Our living standard has been increasing ever since. We have made enough money from growing tea that our children could afford to get married. Our children's lives are getting better and better. Thank you very much.

## References

Agard, Frederick B., 1984, *A course in Romance linguistics*, vol. 2. Washington, DC: Georgetown University Press.

Blair, Frank, 1990, *Survey on a shoestring: a manual for small-scale language surveys.* Dallas, TX.: Summer Institute of Linguistics.

Casad, Eugene H., 1974, *Dialect intelligibility testing.* Dallas, TX.: Summer Institute of Linguistics.

Chiang, Wen-yu, 1996, *The intonation of interrogative sentences in Nantou Atayal.* National Science Council Report NSC 83-0301-H002-072.

Dahl, Otto Christian, 1976, *Proto-Austronesian.* London: Curzon Press. (2nd, revised edition; 1st edition, 1973, Oslo)

Egerod, Søren, 1966, A statement on Atayal phonology. *Artibus Asiae Supplementum* 23(1):120–130. *Felicitation volume for the seventy-fifth birthday of Professor G.H. Luce.*

Frantz, Donald G., 1970, A PL/1 program to assist the comparative linguist. *Communications of the Association for Computing Machinery* 13(6):353–356.

Grimes, Joseph E., 1985, Correlations between vocabulary similarity and intelligibility. *Notes on Linguistics* 41:19–33. Reprinted in Eugene H. Casad, ed. 1992, *Windows on bilingualism*, 17–32. Dallas, TX.: Summer Institute of Linguistics.

—— 1995, *Language survey reference guide.* Dallas, TX.: Summer Institute of Linguistics.

Halim, Amran, Lois Carrington and Stephen A. Wurm, eds, 1982, *Papers from the Third International Conference on Austronesian Linguistics,* vol. 2: *Tracking the travellers.* Canberra: Pacific Linguistics.

Hirano, Takanori, 1972, A study of Atayal phonology. MA thesis, Kyushu University, Japan.

Li, Paul Jen-kuei, 1980, The phonological rules of Atayal dialects. *BIHP* 51(2):349–405.

—— 1981, Reconstruction of Proto-Atayalic phonology. *BIHP* 52(2):235–301.

—— 1982a, Linguistic variations of different age groups in the Atayalic dialects. In Chang Kun et al., eds *Studies in linguistics presented to Dr. Fang-kuei Li on his eightieth birthday*, 167–191. Tsing Hua Journal of Chinese Studies, New Series 14.

—— 1982b, Atayalic final voiced stops. In Halim, Carrington and Wurm, 1982:171–185.

—— 1985, Linguistic criteria for classifying the Atayalic dialect groups. *BIHP* 56(4):699–718.

—— 1993, The distribution and movement of the Austronesian languages on Taiwan [in Chinese]. *Proceedings of the First International Symposium on Languages in Taiwan.*

—— 1996, *The Formosan tribes and languages in I-Lan* [in Chinese]. Monograph Series of I-Lan History, Linguistics 1. Yilan: County Government.

Milliken, Margaret E., 1988, Phonological divergence and intelligibility: a case study of English and Scots. PhD dissertation, Cornell University.

Milliken, Margaret E. and Stuart R. Milliken, 1996, System relationships in assessing dialect intelligibility. *Notes on Linguistics* 72:15–31.

Starosta, Stanley, Andrew K. Pawley and Lawrence A. Reid, 1982, The evolution of focus in Austronesian. In Halim, Carrington and Wurm, eds 1982:145–170.

Tsuchida, Shigeru, 1980, Linguistic position of Sikikun and Mnawyan: linguistic bases of subgrouping Squlyeq and Ts'ole dialects in Atayal. Handout for a talk given at Academia Sinica.

Tu, Wen-Chiu, 1994, A synchronic classification of Rukai dialects in Taiwan: a quantitative study of Mutual intelligibility. PhD dissertation, University of Illinois at Urbana-Champaign.

Wimbish, John, 1989, *WORDSURV: a program for analyzing language survey word lists.* Dallas, TX.: Summer Institute of Linguistics.

Yamada, Yukuhiro and Ying-chu Liao, 1974, A phonology of Tayal. *Research Reports of Kochi University* 23(6):1–9.

Yang, Hsiu-fang, 1976, The phonological structure of the Paran dialect of Sediq. *BIHP* 47(4):611–706.

# 5 *Notes on the prehistory and internal subgrouping of Malayic*

MALCOLM D. ROSS

## 1 Introduction[1]

Adelaar (1992a) provides us with a reconstruction of the phonology and parts of the morphology of Proto Malayic, the language ancestral to the Malayic group proposed by Hudson (1970). The reconstruction is made on the basis of Standard Malay, Melayu Betawi (Jakarta Malay), Minangkabau, Seraway, Banjar Hulu and Iban, with less systematic reference to other communalects and to Old Malay, but the assumption is that Proto Malayic was also ancestral to a whole range of communalects scattered across Indonesia and Malaysia.[2] Adelaar (1992a) is a revision of the author's PhD dissertation (Adelaar 1985), and since the latter's appearance there has been a measure of controversy as to which languages should be included within the Malayic group and as to whether the name 'Malayic' should be used as Adelaar has used it or reserved for a larger grouping (Nothofer 1988; Blust 1988; Adelaar 1991). I shall not address either of these questions here, except with regard to the position of Old Malay. Controversy about the membership of the Malayic subgroup has generally concerned the inclusion of additional languages, not the exclusion of any of Adelaar's member languages,[3] and if one accepts the phonological innovations which Adelaar uses to define Malayic (and hopefully the morphological innovations discussed below), then the admission of further languages to the group should be a relatively

1 This paper is an outcome of research conducted during a visiting professorship in the Department of Southeast Asian Studies at the J.W. Goethe-Universität Frankfurt during the winter semester of 1998–99. I would like to thank Bernd Nothofer, who made this research possible, and the Deutsche Forschungsgemeinschaft for financial support.

2 Since Adelaar (1992a) was written in the early 1980s there has been increased interest in spoken varieties of Malay across both Malaysia and Indonesia, some of which have apparently undergone simplification as a result of use in contact situations. Many of these show little evidence of the bound morphemes discussed in this paper (although most reflect **di-*), and I infer that these morphemes have been lost. Of the communalects used by Adelaar, only Melayu Betawi is a contact variety. The others have probably been transmitted more or less continuously from one generation to the next, and therefore should provide reliable evidence for reconstruction.

3 One may note in passing that if morphology is taken to have the value attributed to it in this paper, then at least one so-called Malay dialect, Macassarese Malay, must be excluded from Malayic since, as Steinhauer (1988) points out, its morphology reveals that it is a relexified form of Macassarese, i.e. Malayicised Macassarese rather than Macassarese Malay.

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 97–109.
Canberra: Pacific Linguistics, 2004.

straightforward matter. Instead, my concern here is to use Adelaar's pioneering work (which is a model of linguistic reconstruction) and to take it a step further by examining the internal subgrouping of his Malayic communalects. The regions where these communalects are spoken are shown in Map 1 (reproduced from Adelaar 1992a).

Adelaar does not propose an internal subgrouping of Malayic.[4] My attempt to do so here is based almost entirely on the material and analysis in Adelaar (1992a), as well as Adelaar's (1992b) article on Salako and his (1999) unpublished morphological sketch of Salako. Furthermore, my observations are limited entirely to innovations in morphology, mostly bound morphology, within the Malayic languages. Other standard means of subgrouping are, of course, the establishment of lexical and of phonological innovations. However, a brief survey of the literature on the languages of the Malayic heartland quickly shows that distinguishing among lexical retentions, innovations and borrowings is a difficult task (e.g. Blust 1981; Nothofer 1985, 1988; Adelaar 1991). And the establishment of shared phonological innovations is itself dependent on one's ability to identify inherited lexemes. The present subgrouping proposals are based on the assumption that bound morphology, and especially productive bound morphology, is somewhat more resistant to borrowing than is the lexicon. I say 'somewhat more', because it is clear (i) that bound morphemes *are* sometimes borrowed between quite closely related languages, and (ii) that massive lexical borrowing can result in the borrowing of morphemes between less closely related languages (cf. English *-ion, -tion*, which entered the language through borrowings from French). I will discuss a case of (i), the passive marker *di-* in Indonesian languages, below.

There are two proposals which I will make here on the basis of morphological innovations:

(1) that Old Malay, despite its conventional name, was not a Malayic communalect;

(2) that the remaining Malayic communalects discussed by Adelaar can be divided into two putative subgroups: a small one consisting of the Western Malayic Dayak communalects which form the dialect chain that includes Salako, Ahe and Belangin (Adelaar 1999), and a large, less well-founded, group consisting of all the other Malayic communalects.[5]

One can legitimately argue that my first proposal is simply a matter of terminology. If Old Malay reflects the phonological innovations listed by Adelaar (1992a:2), then, one may argue, it is a Malayic communalect. However, as far as I can determine, it has never been demonstrated that it does reflect them. In any case, since all Adelaar's Malayic communalects other than Old Malay also share two morphological innovations, we need a name for the subgroup which is defined by both these innovations and the phonological innovations, and I find it convenient to use 'Malayic' for this subgroup.

---

4 In Adelaar (1985) he had divided Malayic into Iban and the rest, but in his 1992a revision he deliberately avoids making a statement about internal subgrouping (Adelaar 1991).

5 I have seen no Belangin data and follow Adelaar (1992b) in including it in a group with Salako and Kendayan.

THAILAND
Urak Lawoi' Malay
Achinese
WEST MALAYSIA
Gayo
Batak
SUMATRA
Minangkabau
Rejang
Middle Malay
( incl. Seraway)
Lampung
Jakartanese
JAVA
Sundanese
Javanese
Madurese
Balinese
Malay
Sasak
BRUNEI
SABAH
( Murut Languages )
MINDANAO
Iban
Brunei Malay
SARAWAK
Selako
Tioman Malay
Kendayan
Mualang
KALIMANTAN
Ngaju
Banjar Hulu
Maanyan
Banjar Kuala
Manado Malay
Bacan Malay
IRIAN JAYA
SULAWESI
Ambon
Ambon Malay
Ujung Pandang
(Macassarese Malay)
Larantuka
Kupang
TIMOR
N
Map of areas where Malayic isolects are spoken

## 2 Reconstructing the Proto Malayic verbal system

Obviously, in order to argue that certain morphological features shared by Malayic communalects and thus attributable to Proto Malayic are indeed innovations, we need to reconstruct the morphology of an interstage or interstages earlier than Proto Malayic, and to this end we need a subgrouping that allows us to identify earlier interstages. The subgrouping hypothesis I adopt here is Blust's, originally presented in 1977, whereby the initial split in the Austronesian family was into the Austronesian languages of Taiwan, which may form several first-order subgroups, and Malayo-Polynesian, the subgroup to which belong all Austronesian languages spoken outside Taiwan.[6] Malayo-Polynesian, in its turn, split into Central/Eastern Malayo-Polynesian and a number of subgroups which are usually labelled 'Western Malayo-Polynesian' but which do not form a higher-order grouping, i.e. there was never a 'Proto Western Malayo-Polynesian' (Ross 1995b). Indeed, we are not yet in a position to put forward any hypothesis according to which Proto Malayic is subgrouped with some other group or groups *within* Malayo-Polynesian. However, it is reasonably clear that there are certain innovations which occurred at some interstage *between* Proto Malayo-Polynesian (PMP) and Proto Malayic, and I label this interstage (or interstages) 'pre–Proto Malayic'.

We can be reasonably certain that PMP had a verbal system of four voices, each of which placed a noun phrase with a particular semantic relationship to the verb — actor, patient, location, conveyance — into the subject slot.[7] Some analysts would say, for various modern languages in Taiwan and the Philippines, that the four voices should be grouped into two: actor (where the subject is the only definite noun phrase) and undergoer (with the variants patient, location, and in some languages conveyance, and where noun phrases other than the subject may also be definite). Whichever analysis we accept, it is clear that in Proto Malayic this system had been reduced to two voices, actor and undergoer, or more conventionally 'active' and 'passive'.[8] Indeed, similar systems are found in Western Malayo-Polynesian languages across Indonesia.

The relevant portion of the PMP system is shown in (1). Like modern Tagalog (Himmelmann 1999), PMP had derivational affixes which formed neutral independent indicative actor-subject verbs (i.e. actor-voice transitives and actor-subject intransitives) from roots. The most important of these affixes were the infix **‹um›*, inserted after the first

---

6 I have summarised the bases of this hypothesis elsewhere (Ross 1992, 1995b) and will not repeat that summary here.

7 The term 'conveyance' is borrowed from Himmelmann (1999) and covers a range of semantic relationships including the thing moved by the action of the verb, the instrument used for that action, and so on. Wolff's (1973) reconstructed four-voice Proto Austronesian system approximates the putative PMP system.

8 Proto Malayic compensated for the reduction of three undergoer voices to one by the innovation of one or two applicative suffixes. The suffix **-i*, which marked the subject of a PMP dependent-clause verb as a location relevant to the action of the verb, became the Proto Malayic locative applicative, marking the subject of a passive verb or the object of an active as a relevant location. The Standard Malay suffix *-kan* marks a passive subject or active object as, among other things, beneficiary, and also performs a variety of other functions. Although there are formally similar suffixes with similar functions in other Malay communalects, Adelaar (1992a) argues, mainly on formal grounds, that a suffix ancestral to *-kan* cannot be reconstructed for Proto Malayic. Although these suffixes are important in the history of the Malay communalects, they will not concern us further in this article.

consonant of the root, and the prefixes **maN-* and **maR-*.[9] The **m-* of **maN-* and **maR-* is historically derived from the infix **‹um›*, and we could alternatively say that PMP had derivational prefixes **paN-* and **paR-* which formed complex stems from which neutral independent indicative actor-subject verbs were then derived by the infix **‹um›*. Stems without **‹um›*, i.e. plain roots and complex stems in **paN-* and **paR-*, occurred in other verb forms (rightmost column of (1)) and in nouns and adjectives. The factors which determined the choice of affix included the morpheme class of the root and probably the semantics of the derivation, but PMP **‹um›* was the default actor subject affix, whilst PMP **paN-* was apparently distributive (marking multiple actions, agents or patients) and **paR-* durative/ intensive.[10]

(1) PMP

| | independent neutral | indicative perfective | projective | dependent or imperative |
|---|---|---|---|---|
| Actor subject | *‹um›√* | *‹um-in›√* | *(‹um›)√-a* | *√* |
| Actor subject | *maN-√* | *naN-√* | ... | *paN-√* |
| Patient subject | *√-en* | *‹in›√* | ... | *√-a* |
| Actor subject | *maR-√* | *naR-√* | ... | *paR-√* |

√ root
... no form has yet been reconstructed for this slot

The forms of undergoer-subject verbs were basically determined by the semantic role of the undergoer: patient, location or conveyance. Only patient forms are shown in (1). As well as the forms discussed in the previous paragraph, PMP also had perfective independent indicative forms used for past events and projective forms used to express intention, possibility and exhortation (Ross 1995a).[11] The perfectives were all historically derived from forms with the infix **‹in›*. Only one projective form, used with an actor subject, is currently reconstructable. This is shown in (1) as **(‹um›)√-a*, the parentheses indicating that **‹um›* may not have been present in PMP (it certainly had been present in Proto Austronesian).

9 The **-N-* of **maN-* combined with root-initial **p, *t, *k* and **c/*s* respectively to give **-m-, *-n-, *-ŋ-* and **-ñ-*, disappeared before a root-initial nasal, and otherwise became a nasal homorganic with the root-initial consonant.

10 This attempt to reconstruct the meanings of PMP affixes was limited to an examination of Ilokano (Rubino 1994), Tagalog (Ramos 1971), the Bisayan dialects (Wolff 1972; Zorc 1977) and Binukid (Post 1992). It is clear that far more work is needed on this aspect of PMP reconstruction.

11 There were also reduplicated forms for the imperfective (Wolff 1973; Reid 1992). My 'projectives' are Wolff's 'subjunctives'.

(2) Pre-Proto Malayic

| | independent indicative — | passive | projective | dependent, imperative, non-verbal |
|---|---|---|---|---|
| Actor subject | *maN-√* | — | *√-a* | *√, paN-√* |
| Patient subject | *A=√* | *‹in›√* | ... | |
| Actor subject | *maR-√* | — | ... | *paR-√* |
| Patient subject | *paR-√* | — | | |

√ root
... no form has yet been reconstructed for this slot
— no form occurred in this slot in the paradigm
A = proclitic agent pronoun

At some pre–Proto Malayic stage, the system underwent a reorientation and reduction, roughly as sketched in (2) — 'roughly' because we do not yet understand enough of the historical development of the Western Malayo-Polynesian languages to know just what happened when, and to what extent similar historical processes occurred independently in different parts of Indonesia and Malaysia. This reorientation process is reconstructed by Wolff (1996). The events which must have occurred are the following, but we cannot be certain of their sequence:

(a) Of the three sets of actor-subject forms, the **maN-√* set seems to have acquired the largest number of members and became the default actor-subject set. The **‹um›√* set was eventually reduced to a few fossils.[12] The **maN-√* set thus became the actor-subject correspondent of the patient-subject set (**√-en* etc.).

(b) At least in the case of the neutral patient-subject form **√-en*, the form of the agent pronoun (e.g. Standard Malay *ku-*) became the crucial marker of a patient-subject form, and the suffix was eroded, leaving **√*, i.e. the same form as in the imperative.

(c) The **maR-√* set acquired a patient-subject form of its own, recruiting the imperative in **paR-√* by analogy with the **maN-√* set, where the patient-subject form was the same as the imperative.

(d) The perfective infixed form **‹in›√* was reinterpreted as a passive, probably an agentless passive or one with a third-person agent, and other perfective forms disappeared.

Stage (d) is inferred on the basis of other Indonesian languages (Old Malay, Javanese, Nias and a number of languages of Sulawesi): it is possible that **‹in›√* simply disappeared along with the other perfectives (see below).

12 Adelaar (1992a:194) reconstructs three pairs in which **m*-initial forms, apparently reflecting earlier **‹um›*, contrast with vowel-initial root forms: **mimpi/*impi* 'dream', **mampus/*ampus* 'wiped out, gone', **minum/*inum* 'drink'. He also reconstructs the triplet **mipis/*nipis/*tipis* 'thin', which is more difficult to explain.

(3) Proto Malayic

| | independent indicative – | independent indicative passive | projective | dependent, imperative, non-verbal |
|---|---|---|---|---|
| Actor subject | *maN-√* | __ | *√-a* | *√, pAN-√* |
| Patient subject | *A=√* | *di=*[ANP]*√* | ... | |
| Actor subject | *bAR-√* | __ | ... | *pAR-√* |
| Patient subject | *pAR-√* | __ | | |

√ root
... no form has yet been reconstructed for this slot
__ no form occurred in this slot in the paradigm
= clitic boundary; *A=* proclitic agent pronoun
ANP agent noun phrase (optional)

The morphemes which Adelaar (1992a) reconstructs for Proto Malayic are set out in (3), with the addition of **di=*[ANP]*√*. This morpheme set differs from (2) in only two respects, and I take these to represent Proto Malayic morphological innovations. These are:

(4) Proto Malayic morphological innovations:
  (i) passive **‹in›√* — if the reconstruction in (2) is correct — is replaced by Proto Malayic **di=*[ANP]*√*;
  (ii) the earlier actor-subject form **maR-√* is replaced by Proto Malayic **bAR-√*.

Before we can proceed to a discussion of these innovations and their implications, however, I need to make a number of points about (3).

In (3) I have retained the same labels to the left as in (1) and (2). I have done this only for the sake of comparison. Conventional Indonesianist terminology would replace my 'Actor subject' and 'Patient subject' with active and passive, with the implication that in (3) 'Patient subject' √ and 'Patient subject' passive **di=*[ANP]*√* formed a single passive paradigm — and indeed they did. Proto Malayic 'Patient subject' √ was evidently used with agent prefixes in the first and second persons, **di=*[ANP]*√* in the third person.

Adelaar uses **A* where PMP **a* and **e* have merged in all witness communalects, here in prepenult syllables. His reconstruction of **√-a* differs formally from mine in that he writes **√-aʔ*, inserting the glottal stop on the basis of its only two Malayic reflexes, in Salako and Kendayan. I omit it because these two communalects insert a glottal stop after certain final vowels. Thus they reflect the Proto Malayic locative applicative verbal suffix **-i* as *-iʔ*. Since non-Malayic witnesses also reflect **√-a*, I infer that neither **-i* nor **-a* had a glottal stop in Proto Malayic.

The Proto Malayic passive **di=*[ANP]*√* is my own addition to Adelaar's reconstructions, but I have reconstructed it on the basis of the material in Adelaar (1992a:161–162) and Adelaar (1992b). As Adelaar points out, in Salako and Kendayan *di-* is procliticised to the verb only if the agent is not expressed, as in the Kendayan sentence in (5) (Kendayan examples are from Thomas et al. 1984):

(5) *Buku singkara di-baca.*
book story PASS-read
'The story book was read.'

If there is a nominal or pronominal agent it intervenes **between** *di-* and the verb, as the formulation **di*=[ANP]√ indicates.[13] In (6) and (7) respectively the agent is a proper name and a common noun phrase:

(6) *Buku singkara di-Adiʔ maca.*
book story PASS-Adik *N*:read
'Adik read the story book.'

(7) *Murid di-guru-nya ngajar.*
pupil PASS-teacher-P:3S *N*:teach
'The pupil was taught by his teacher.'

In (8) the agent is a pronoun:[14]

(8) *Bini mudaʔ-nya udah di-nya nyare-aʔn.*
wife young-P:3S PF PASS-P:3S *N*:divorce-APPL
'He has already divorced his young wife.'

I take this to have been the Proto Malayic construction, but with one difference. In Proto Malayic, **N-* marked an actor subject. The presence of *N-* in the examples above is an innovation, and I return to this below. There has been ample discussion in the literature of the origins of passive *di-* in Malay communalects and a good many other Indonesian languages, and three hypotheses have been put forward:

A. It is derived from the agentive preposition **di* (Aichele 1942–43).

B. It reflects the third person post–Proto Malayic **d-ia* (i.e. Proto Malayic *ia* with accreted topic-marking preposition **di*): this is compatible with the fact that it seems first only to have occurred with third person agents (Adelaar 1992a:162).

C. It reflects passive **ni-*, itself derived from pre–Proto Malayic **‹in›*, reconstructed in (2) above. The prefix **ni-* underwent an idiosyncratic sound change (obstruent denasalisation) by which also *ber-* is derived from **mAR-* (De Casparis 1956:24; Teeuw 1959:141–144).

Adelaar (1992a:161–162) points out that *di-* is the only prefix not to have undergone prepenultimate neutralisation, implying it was not a prefix when neutralisation occurred. This observation lends support to A and B, but not C. He in any case suggests (1992a:163) that **bAR-* developed from earlier **mAR-* via consonant epenthesis: **mAR-* was most often prepenultimate and therefore unstressed, giving [mr-] ~ [mbr-], then with denasalisation [br-] and finally with schwa epenthesis [bər-].[15] This explanation will not work for **ni- > di-*, and there is no obvious ground to favour C. I agree with Adelaar (1992a:163; 1999) that the evidence from Salako and Kendayan favours A, and this is also supported by the fact that *di* is an agentive preposition in Minangkabau. However, Adelaar does not reconstruct *di* as

---

13 Adelaar (1999) notes that in Salako an actor noun phrase immediately preceding the verb is not necessarily preceded by *di-*.

14 Adelaar also reports an example with a non-third person pronoun: *di-ku-kuruk*$^{a}$ 'caged by me'.

15 If this explanation is correct, it implies that **bAR-* was phonetically [bər-] and that Proto Malayic **bAR-* can be replaced by **beR-*.

a passive marker, partly on the grounds that the Old Malay inscriptions have *ni-* rather than *di-*. I return to this matter below.

The insertion of **bAR-* in an Actor subject slot in (3) is my decision, not Adelaar's. He examines the history of prefixes reflecting **bAR-* in an earlier (1984) article, and concludes that the function of **bAR-* was to form intransitives, and that it had a paradigmatic relationship to **pAR-* which formed transitives. The reason for his conclusion is that most reflexes of **bAR-* in present-day Malayic communalects do indeed form intransitives. It seems to me, however, that there are reasons to infer that its Proto Malayic function was to form a class of transitives, and that **bAR-* was the Actor subject correspondent of Patient subject **pAR-*.

The first reason is that, as Adelaar says, **bAR-* reflects PMP **maR-*. There is no doubt that **maR-* was an Actor subject prefix, evidently corresponding to the Patient subject transitive prefix **pAR-*.

Secondly, Iban *be-/b-/beC-/ba-*, reflecting **bAR-*, is functionally equivalent to Standard Malay *ber-*, i.e. it occurs in intransitives, but it may also co-occur with *-ka* to produce a transitive. Where this transitive is imperative or passivised, *be-* is replaced by *pe-*: e.g. *be-jalay* 'walk', *be-jalay-ka* 'move something', *pe-jalay-ka* 'move it!', *di-pe-jalay-ka* 'be moved'. This seems to reflect an earlier Actor subject/Patient subject transitive relationship.

Thirdly, Roolvink (1965) examines *ber-* and *per-* in Classical Malay texts, and finds that *ber-* marked actor-subject verbs, both intransitive and transitive, until the end of the 18th century, whilst *per-* marked patient-subject transitives. Roolvink thinks that *ber-* lost its transitive meaning through competition with *meN-*, resulting in the innovation of *memper-* as the active equivalent of passive *per-*. That *memper-* is a relatively recent formation is not in doubt, as the sequence *meN-per-* has no cognates outside Malay dialects and is in any case morphologically irregular, in that nasal assimilation should give rise to **memer-*, but it does not.

Since (i) Proto Malayic **bAr-* was descended from the PMP Actor subject prefix **maR-*, which could be used both intransitively and transitively, and (ii) Roolvink shows that *ber-* had these functions in 18th century Classical Malay, it is a reasonable inference that Proto Malayic **bAr-* was also used both intransitively and transitively.

## 3 The Proto Malayic passive and the position of Old Malay

It will perhaps by now be evident to the reader why I propose that Old Malay was not a Malayic communalect: it reflects neither of the morphological innovations in (4).

If we accept passive **di=*[ANP]√ as the Proto Malayic passive construction, this has certain further implications. First, if **di-* is indeed *not* derived from pre–Proto Malayic **‹in›*, then there is no pressing reason to reconstruct **‹in›* in (2), and this is why I questioned its reconstruction above. It may be that **‹in›* had disappeared as a passive marker and that the plain stem had taken over this function considerably earlier than the innovation of **di=*[ANP]√. It may even be (but this seems less likely) that **‹in›* never underwent the change from perfective to passive marker in a precursor of Proto Malayic, but simply disappeared along with the other markers of the perfective reconstructed in (1).

The second implication is this. If we reconstruct **di=*[ANP]√ as the Proto Malayic passive construction, where '...' represents the agent noun phrase, then we have a source for the passive prefix *di-* in Malayic communalects other than those of Salako and Kendayan. In

the latter, the preposition *di-* is only partially grammaticised as a passive prefix with agentless passives, as in (5). In the other communalects, this process has been completed: *di-* is fully grammaticised as a prefix, and the agent noun phrase follows the verb. An important implication here is that this grammaticisation process is not reversible. Salako and Kendayan reflect an early stage in the development of the Malayic passive, not a late one, as it is not credible to reconstruct the opposite developmental path whereby the agent noun phrase comes to be inserted between an already existent prefix and the verb.

This in turn helps us to approach another problem. The passive marker *di-* occurs in languages scattered across Indonesia, e.g. Lampung, Batak, Sundanese, modern Javanese, and in some South Sulawesi languages. Most of these are not candidates for inclusion in a 'greater Malayic' grouping, and we must therefore infer either (i) that **di-* is very old, i.e. that it occurred in a language ancestral to all modern languages that have *di-*, or (ii) that it has been borrowed from language to language. The very fact that we can trace the development of *di-* in Malayic speaks clearly against (i) and suggests that Malay is the source of the borrowing proposed in (ii). Philological evidence from Javanese also speaks against (i), as we know that Javanese *di-* is a borrowing (the inherited Javanese passive marker is *‹in›*). Why *di-* should have been so readily borrowed is not clear to me. Strømme (1994:94–95) points out that the inherited South Sulawesi passive marker *ni-* is homophonous with the first person plural exclusive ergative pronominal prefix *ni-* and that they occupy the same slot with patient-subject verbs. They are therefore open to confusion. She also reports that younger speakers of Mamuju are currently switching from inherited passive *ni-* to Malay *di-*. How widespread this causation may be, I do not know, but it would not account for the Javanese borrowing.

The observations in the previous paragraph raise another issue. John Bowden (pers. comm.) points out that *di-* often occurs in Malay communalects that otherwise have little affixation. This implies that in such communalects the original reflex of Proto Malayic **di-* may have been lost along with much of the verbal morphology and that the present reflex of **di-* is also the outcome of borrowing. If *di-* were the outcome of borrowing in the communalects on which Adelaar's (1992a) reconstruction is based, then the reconstruction of **di-* in Proto Malayic could be called into question (although it would have to be reconstructed at some interstage). However, of these communalects only Melayu Betawi displays obvious indicators of contact-induced change or morphological borrowing, so this is probably not a problem.

## 4 The internal subgrouping of Malayic

We come finally to my other subgrouping proposal: that the Malayic communalects discussed by Adelaar, other than Old Malay, can be divided into two subgroups: a small one consisting of the communalects of the Western Malayic Dayak communalects, including Salako, Ahe, Kendayan and Belangin, and a large one consisting of all the other Malayic communalects. I will call the large subgroup 'Nuclear Malayic'. The innovations which define Nuclear Malayic are:

(9) Nuclear Malayic innovations:

(a) as noted above, grammaticisation of *di-* as passive marker is complete in Nuclear Malayic, but not in Western Malayic Dayak.

(b) PMP evidently retained the Proto Austronesian pattern whereby, in a narrative sequence, verbs in clauses after the first took the atemporal form (Ross 1995a). Adelaar (1992a:161) reports that Salako and Kendayan use the bare stem for the atemporal, reflecting the PMP pattern, whereas Nuclear Malayic communalects have innovated, using an inflected verb instead of the bare stem.

(c) Salako and Kendayan retain the Proto Austronesian/PMP/Proto Malayic projective marker *-*a* (Adelaar 1992a:163–164) in its reconstructable Proto Austronesian function (Ross 1995a), illustrated in (10) below. Nuclear Malayic has lost it.[16]

(10) Kendayan:

*Ne?,* *kami* *pula?ng-a?.*
grandmother D:1EP return.home-INTENT
'Grandmother, we want to go home.'

*Ampus-a?* *ge??*
go-INTENT Q
'Do you want to go?'

*N-yabarang-a?* *ge?* *ina??*
*N*-cross.over-INTENT Q NEG
'Do you want to cross over or not?'

Western Malayic Dayak, on the other hand, is defined by only one innovation, but a significant one:

(11) Western Malayic Dayak innovation:

(a) According to Adelaar (1992a:161), Salako and Kendayan both use *N*- to mark realis in main clauses, whereas other Malayic communalects retain the PMP pattern whereby **N*- or **maN*- marks a verb as having an actor subject. The use of *N*- in Kendayan passives is seen in (6), (7) and (8) above. Adelaar (1999) describes the use of *N*- in Salako a little differently: *N*- marks transitive verbs and co-occurs both with actor subjects and with verbs that refer to a completed event and have an undergoer subject.

The effect of this subrouping is to split the putative Malayic Dayak group in half, including the eastern languages in Nuclear Malayic and placing the western languages in a subgroup of their own. However, there is nothing especially surprising in this, as the term 'Malayic Dayak' was invented by Hudson (1970) as a label for non-Muslim Bornean Malayic languages, not for a linguistically defined subgroup. Moreover, the proposal that Western Malayic Dayak is a first-order subgroup of Malayic accords with Adelaar's (1992b) observation that they are among the most conservative of Malayic communalects and, more specifically, with Nothofer's (1988) proposal that Salako not be placed under the same node in the family tree as other Malayic communalects, but one node higher. I must admit that the defining innovations of the Nuclear Malayic subgroup are none too strong. Innovations (9b)

16 Old Malay also reflects **-a*, but this is irrelevant since we have excluded it from Malayic. Like its other retentions, it simply suggests that Old Malay belongs to a subgroup more conservative than Malayic, a fact which is not surprising, given the antiquity of the inscriptions.

and (9c) could have been effected independently in different dialects or by contact. Innovation (9a), the reanalysis of a preposition as a passive marker, is less likely to have occurred independently, but, as I noted above, it could readily have spread by contact.

If it is accepted that the homeland of a protolanguage is likely to be somewhere near the seam between its primary subgroups, then this places the Malayic homeland in southwest Borneo, where Western Malayic Dayak and Nuclear Malayic languages are more or less contiguous. This agrees with the proposals of other scholars (Adelaar 1988, 1995; Blust 1988) about the location of the homeland.

## References

Adelaar, K. Alexander, 1984, Some Proto-Malayic affixes. *Bijdragen tot de taal-, land- en volkenkunde,* 140:402–421.

—— 1985, Proto-Malayic: the reconstruction of its phonology and parts of its lexicon and morphology. PhD dissertation, University of Leiden.

—— 1988, More on Proto-Malayic. In Ahmad and Zain, 1988:59–77.

—— 1991, On the internal classification of the Malayic subgroup. Paper presented to the Fifth International Conference on Austronesian Linguistics, Honolulu.

—— 1992a, Proto-Malayic: the reconstruction of its phonology and parts of its lexicon and morphology. Canberra: Pacific Linguistics.

—— 1992b, The relevance of Salako for Proto-Malayic and for Old Malay epigraphy. *Bijdragen tot de taal-, land- en volkenkunde* 148:381–408.

—— 1995, Borneo as a cross-roads for comparative Austronesian linguistics. In Peter Bellwood, James Fox and Darrell Tryon, ed. *The Austronesians: historical and comparative perspectives.* Canberra: Department of Anthropology, Research School of Pacific and Asian Studies, The Australian National University.

—— 1999, Salako morphology and the interrelation between voice, mood and aspect. Unpublished MS.

Ahmad, M.T. and Z.M. Zain, 1988, *Rekonstruksi dan cabang-cabang ahasa Melayu induk.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Aichele, W., 1942–43, Die altmalaische Literatursprache und ihr Einfluss auf das Altjavanische. *Zeitschrift für Eingeborenen-Sprachen* 33:37–66.

Blust, Robert A., 1977, The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report, *Working Papers in Linguistics of the University of Hawaii,* 9(2):1–15.

—— 1981, The reconstruction of Proto-Malayo-Javanic: an appreciation. *Bijdragen tot de taal-, land- en volkenkunde* 137:456–469.

—— 1988, Malay historical linguistics: a progress report. In Ahmad and Zain, 1988:1–33.

Casparis, J.G. de, 1956, *Selected inscriptions from the 7th to the 9th century A.D. Prasasti Indonesia II.* Bandung: Masa Baru.

Himmelmann, Nikolaus P., 1999, Lexical categories and voice in Tagalog. Unpublished MS.

Hudson, Alfred B., 1970, A note on Selako: Malayic Dayak and Land Dayak languages in west Borneo. *Sarawak Museum Journal* 18:301–318.

Nothofer, Bernd, 1985, The subgrouping of the languages of the Javo-Sumatra Hesion: a reconsideration. *Bijdragen tot de taal-, land- en volkenkunde* 141:288–302.

—— 1988, A discussion of two Austronesian subgroups: Proto-Malay and Proto-Malayic. In Ahmad and Zain,1988:34–58.

Post, Ursula, 1992, *Binukid dictionary. Studies in Philippine Linguistics* 9/2. Manila: Linguistic Society of the Philippines and Summer Institute of Linguistics.

Ramos, Teresita V., 1971, *Tagalog dictionary*. Honolulu: The University Press of Hawaii.

Reid, Lawrence A., 1992, On the development of the aspect system in some Philippine languages. *Oceanic Linguistics* 31:65–91.

Roolvink, R. 1965, The passive-active *per-/ber-//per-/memper-* correspondence in Malay. *Lingua* 15:310–337.

Ross, Malcolm D., 1992, The sound of Proto-Austronesian: an outsider's view of the Formosan evidence. *Oceanic Linguistics* 31:23–64.

—— 1995a, Reconstructing Proto Austronesian verbal morphology: evidence from Taiwan. In Paul Jen-kuei Li, Dah-an Ho, Ying-kuei Huang and Cheng-hwa Tsang, ed. *Austronesian studies relating to Taiwan*, 727–791. *Symposium Series of the Institute of History and Philology, Academia Sinica* 4. Taipei: Institute of History and Philology, Academia Sinica.

—— 1995b, Some current issues in Austronesian linguistics. In D.T. Tryon, ed. *Comparative Austronesian Dictionary* 1, 45–120. Berlin: Mouton de Gruyter.

Rubino, Carl R.G., 1994, A root based dictionary of the Ilocano language. Unpublished MS.

Steinhauer, Hein, 1988, Malay in east Indonesia: the case of Macassarese Malay. In Ahmad and Zain 1988:108–151.

Strømme, Kari K., 1994, Person marking in the Mamuju language. In René van den Berg, ed. *Studies in Sulawesi linguistics*, part 3. *NUSA* 36:91–113. Jakarta: Badan Penyelenggara Seri NUSA.

Teeuw, A., 1959, The history of the Malay language. *Bijdragen tot de taal-, land- en volkenkunde* 115:138–156.

Thomas, Yoseph, Hery Suyatman, J.B. Mangunsudarsono and Rusmani Handayani, 1984, *Morfologi dan sintaksis Bahasa Kendayan*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Wolff, John U., 1972, *A dictionary of Cebuano Visayan*, 2. Ithaca N.Y. and Manila: Cornell University Southeast Asia Program and Linguistic Society of the Philippines.

—— 1973, Verbal inflection in Proto-Austronesian. In Andrew Gonzalez, ed. *Parangal kay Cecilio Lopez*, 71–91. Quezon City: Linguistic Society of the Philippines.

—— 1996, The development of the passive verb with pronominal prefix in western Austronesian languages. In Bernd Nothofer, ed. *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, 15-40. Hamburg: Abera.

Zorc, R. David, 1977, *The Bisayan dialects of the Philippines: subgrouping and reconstruction*. Canberra: Pacific Linguistics.

# 6 *Internal subgrouping and pronominal paradigmaticity: the case of Nuclear Micronesian*

JAE JUNG SONG

## 1 Introduction[1]

There are at least two tree models available in the literature which depict the internal genetic relationships of the Nuclear Micronesian languages: (i) what Rehg (1995:311) calls the flat tree model; and (ii) the stratified tree model proposed by Jackson (1983:433). The former, foreshadowed in Bender (1971), is reproduced in Figure 1, and the latter in Figure 2. Note that the broken lines in the stratified tree model indicate where 'the historical relationships are not reasonably clear' (Jackson 1983:433).[2]

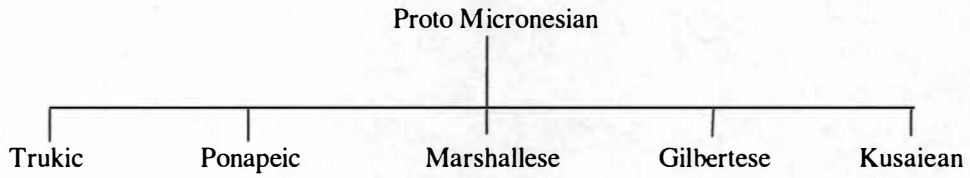



**Figure 1:** A flat tree model of the Nuclear languages of Micronesia

1 I am indebted to Barry Blake and John Bowden for their most useful comments on earlier drafts of this paper. The abbreviations used are: Kir = Gilbertese, Ksr = Kusaiean, Map = Mapian, Mc = Micronesian, Mok = Mokilese, Mrs = Marshallese, Mrt = Mortlockese, PCMc = Proto Central Micronesian, PCTk = Proto Central Trukic, PEO = Proto Eastern Oceanic, PETk= Proto Eastern Trukic, PMc = Proto Micronesian, PNTk = Proto Nuclear Trukic, POc = Proto Oceanic, Pon = Ponapean, PPp = Proto Ponapeic, PSTk = Proto Sonsorol-Trukic, PTk = Proto Trukic, PTk-Pp = Proto Trukic-Ponapeic, Pua = Pulo Annian, Pul = Puluwatese, PWMc = Proto Western Micronesian, PWTk = Proto Western Trukic, Sns = Sonsorolese, Stw-Crl = Satawalese-Carolinian, Trk = Trukese, Uli = Ulithian, Wol = Woleaian.

2 Jackson (1986:214) proposes a stratified tree model which seems to be less circumspect than the one in Figure 2 with respect to the position of Ulithian (and probably Pulo Annian), and Ponapeic (or his PPp).

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 111–128.
Canberra: Pacific Linguistics, 2004.

The flat tree model, in which five subgroups are recognised, is also adopted by Bender and Wang (1985:80), and Rehg and Bender (1990:2), pending further evidence for higher-level subgroups. The received view in Micronesian linguistics seems to be that the flat tree model may not be correct, especially in the light of the evidence that Jackson (1983; 1986) adduces against it (e.g. Rehg 1995:311).[3] The stratified tree model in Figure 2, on the other hand, has not yet been accepted in full by Micronesian specialists; Rehg and Bender (1990:24), for instance, point out cautiously that '[w]hether the other languages are coordinate with [Trukic and Ponapeic], or whether higher-level subgroups [as represented in Figure 2] exist within Micronesian is less certain [than whether Trukic and Ponapeic are well-defined subgroups within Micronesian]' (also see Rehg 1995:317 for a similar point of view).

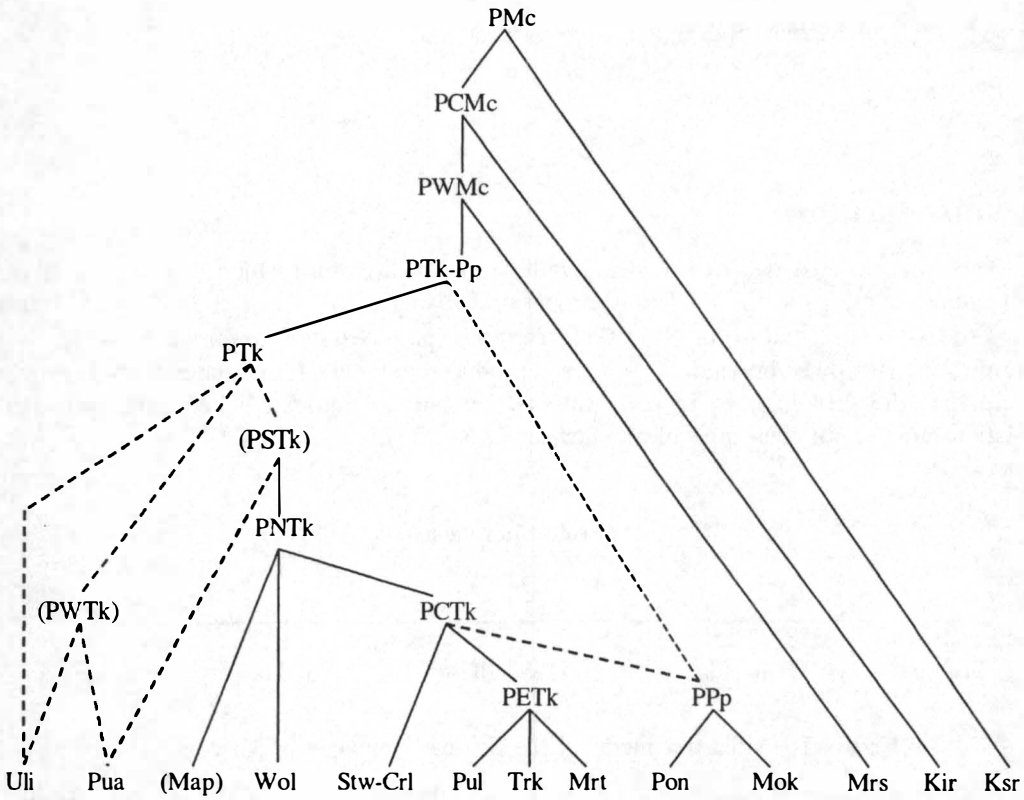



**Figure 2:** Hisorical relationships within the Micronesian subgrouping of Oceanic

In common with others who previously commented on the family tree theory (e.g. Bloomfield 1933:311–318; Southworth 1964; Grace 1986:1), Rehg (1995:313) identifies the 'use of branching trees to depict linguistic relationships' as the major drawback of such tree models as those in Figures 1 and 2, because it is totally implausible to accept 'the uniform parent languages and their sudden and clear-cut splittings as historical realities' (Bloomfield 1933:311). Apart from the question as to whether PMc was a completely uniform parent language, 'if we interpret [Jackson's tree] as a literal model of migration patterns, then we must conclude that Micronesia was settled by a series of discrete moves

3 Rehg (1995:311) also points out that, if the flat tree model is wrong, the explanation based on geography of the patterns of Nuclear Micronesian languages (e.g. Irwin 1992) is mistaken.

through the islands, and that at each point where we identify a subgroup, there was a pause of sufficient duration to allow a unique set of innovations to develop by which we identify the subgroup' (Rehg 1995:314). Rehg (1995:314) thinks that this is an unlikely scenario of the settlement of the entire region of Micronesia. As evidence for this, he points to Jackson's (1983:413–431) work on the Trukic/Ponapeic subgroup(s). Being well-defined subgroups within Micronesian by virtue of sharing a substantial number of unique innovations, Trukic and Ponapeic prove to be problematic in that there are several major innovations which are not due to drift or independent development, but are uniquely shared only by Ponapeic and the Central Trukic languages, e.g. spirantisation and loss of PMc **t*. This is why PPp is connected to both PTk-Pp and PCTk by broken lines in Figure 2 (see Jackson 1983:421–428 for detailed discussion).

These fundamental problems with the family tree model notwithstanding, Rehg (1995: 317–318) is of the opinion that '[t]here are ... circumstances under which family trees can be employed without distorting historical facts ... [w]hen speech communities divide from each other at a single point in time, and when they remain relatively or totally isolated, or when, if contact occurs, its effects can be discerned' (see e.g. Rehg and Bender 1990 for such contact-induced effects in Mokilese). Rehg (1995:317–318) believes that most, if not all, higher levels of linguistic relationships within Micronesia may be of this type. He also points to the insightful works of Robert Blust, Andrew Pawley, and Malcolm Ross, who all make use of trees when and where appropriate.[4]

But, as Rehg (1995:318) himself asks, the question may then be: 'How in our research do we determine when trees are appropriate?' Part of the answer is, Rehg (1995:318) suggests, that the distribution of all innovations must first meticulously be tracked 'without regard to preconceived notions of language and subgrouping boundaries', and can then perhaps be compared with available (or competing) tree models.

Rehg's suggestion is taken here to be a call for papers to identify innovations or properties which can be utilised for an understanding of the internal genetic relationships of the Nuclear Micronesian languages. The primary purpose of the present chapter is to identify and document one such property: the paradigmaticity of the focus and possessive pronoun systems.

## 2 Paradigmaticity as probative evidence

Nichols (1996) argues convincingly that demonstration of genetic relationships among languages through systematic correspondences in vocabulary is not the operating procedure for the application of the comparative method. In fact, such demonstration can only be carried out by adducing 'evidence [that] is primarily grammatical and includes morphological material with complex paradigmatic and syntagmatic organization' (Nichols 1996:41). For example, the segment of adjectival morphology of Latin and Greek in Table 1 is regarded as such evidence in that it has what Nichols (1996:46) calls 'multidimensional paradigmaticity'.

---

4 Ross (1988:9–11), in an attempt to make a distinction between language separation and dialect differentiation in genetic trees, adopts both standard branching nodes, and innovative double horizontal lines, the latter intended to capture dialectal linkages. Rehg (1995:317) indeed makes use of Ross's double horizontal lines in order to represent in his genetic tree such problematic innovations as the spirantisation and loss of PMc **t* discussed earlier.

**Table 1:** Partial adjectival morphology of Latin and Greek

| | Masculine | Feminine | Neuter |
|---|---|---|---|
| **Latin:** | | | |
| Nominative | *-us* | *-a* | *-um* |
| Accusative | *-um* | *-am* | *-um* |
| **Greek:** | | | |
| Nominative | *-os* | *(*)-ā* | *-on* |
| Accusative | *-on* | *(*)-ān* | *-on* |

Table 1 involves two dimensions of paradigmaticity: (i) case (nominative and accusative); and (ii) gender (masculine, feminine, and neuter). Number, if also included in Table 1, would be a third dimension. Moreover, in both Latin and Greek the masculine and neuter adjectival endings are identical to **o*-stem noun endings, and the feminine adjectival endings to **a*-stem noun endings. This can be taken to be a fourth dimension of paradigmaticity (Nichols 1996:46). There are, in addition to the abstract paradigmaticity, phonologically specific or concrete fillers (or forms) and grammatically specific (or designated) functions for the slots in the paradigm in Table 1. This entire system with multiple paradigmaticity and a degree of phonological and functional specificity can thus be understood to constitute a piece of probative evidence for the genetic relatedness of Latin and Greek.[5]

At first glance, 'personal pronouns offer a good example of a systematically structured and phonologically filled lexical field' (Nichols 1996:54), because they may involve persons (first, second, and third), numbers (singular, dual, plural etc.), functions (focus, subject, object, possessive etc.) and even genders (animate, inanimate, human, nonhuman etc.). But Nichols (1996:54) hastens to sound a warning that they do not constitute probative evidence for genetic relatedness, because 'the forms of first and second persons, and of singular and plural numbers, are not independent; that is, in a personal pronoun system the relation of paradigmaticity to coding phonological form is nonarbitrary'. In other words, personal pronouns are very likely to exhibit 'their paradigmatic relationships and their deictic semantics' by means of consonant symbolism or 'phonosymbolism'.[6] Thus, 'the presence of a nasal in at least one of the personal pronoun forms is to be expected and the presence of a labial in one of the forms makes it quite likely that the other person or number form (or both) will contain a dental' (Nichols 1996:54). For this reason, personal pronouns may not freely be accepted as probative evidence of genetic relatedness (also see Meillet 1958:89–90).[7]

---

5 For instance, Meillet (1958:91, 97) is quoted by Nichols (1996:47) as saying (Nichols's own translation of both quotations):

> Grammatical correspondences are proof, and rigorous proof, provided one makes use of the material detail of the forms and that it is established that particular grammatical forms used in the languages under consideration go back to a common source.
>
> While one can initially establish vocabulary resemblances between two or several languages as an indication of where to do further research, this cannot furnish a definitive demonstration; vocabulary can only orient the research, and proof comes from elsewhere.

6 In the present chapter, I will not be concerned with determining whether or not phonosymbolism really is inherent in personal pronouns. Nonetheless, one cannot be too careful to be mindful of such phonosymbolism.

7 In order to strengthen her argument in support of phonosymbolism in personal pronouns, Nichols (1996:56) also quotes Meillet (1958:89–90) as saying (Nichols's translation):

Personal pronoun systems, however, can contribute to the establishing of internal genetic relationships where phonosymbolism can be kept at bay, as it were. Suppose a given group of languages are taken on the basis of other evidence to form a genetic group, but their internal-subgrouping details are yet to be worked out or are not well understood. Under these circumstances, since it is established at least that these languages all come from a single source, there is no need to be overly wary of phonosymbolism in their personal pronoun systems. The languages in question are expected to have more or less inherited the pronoun system of their parent language. If form/function relationships in the personal pronoun system of any of these languages cannot be traced back to the parent language (i.e. abstract paradigmaticity, phonologically specific fillers, and grammatically specific functions), these can then be analysed as innovations for purposes of internal subgrouping. In other words, the phonosymbolism residual in personal pronoun systems can be ignored in the context of an already established genetic group, and the paradigmaticity of personal pronoun systems can reliably be employed for purposes of internal subgrouping. Indeed, linguists (e.g. Blake 1989, 1990; Ross 1996) have successfully made use of personal pronouns in carrying out comparative work without being too much concerned about phonosymbolism in personal pronoun systems.

With phonosymbolism being 'controlled' in this way, one can proceed to establish systematic form/function correspondences that may be embodied in the equivalent personal pronoun systems across the languages in question. This can be called an 'intrasystemic' comparison. For instance, the pronoun system X (e.g. third person singular) in Language A will be compared with the corresponding pronoun system X (e.g. third person singular) in Language B in terms of both form and function, and so on, as schematised in Figure 3, where double-headed arrows represent the 'loci' of comparison. The pronoun system X here can be one of the pronoun systems that may exist in both Language A and Language B, e.g. focus, subject, object, or possessive.

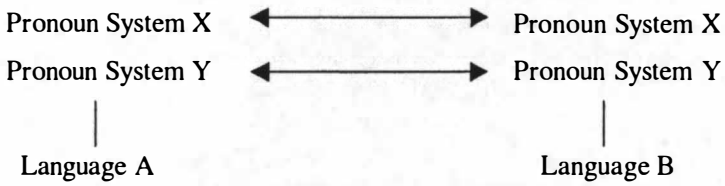



**Figure 3:** Intrasystemic comparison

Systematic form/function correspondences cannot only be sought in the equivalent personal pronoun systems across different languages as in an intrasystemic comparison, but also across the different personal pronoun systems in one and the same language. In what may be called an 'intersystemic' comparison, the different personal pronoun systems in a single language will be compared. For example, the focus pronoun system of Language A

---

It goes without saying that in order to establish genetic relatedness of languages one must disregard everything that can be explained by general conditions common to all languages. For instance, pronouns must be short words, clearly composed of easily pronounced sounds, generally without consonant clusters. The consequence is that pronouns are similar in almost all languages, though this does not imply a common origin. On the other hand, pronouns often show little resemblance in languages that are otherwise quite similar [...] Therefore, pronouns must be used with caution in establishing relatedness of languages.

will be compared with the subject pronoun system of the same language in terms of both form and function, and so on, as schematised in Figure 4 (where double-headed arrows represent the 'loci' of comparison).

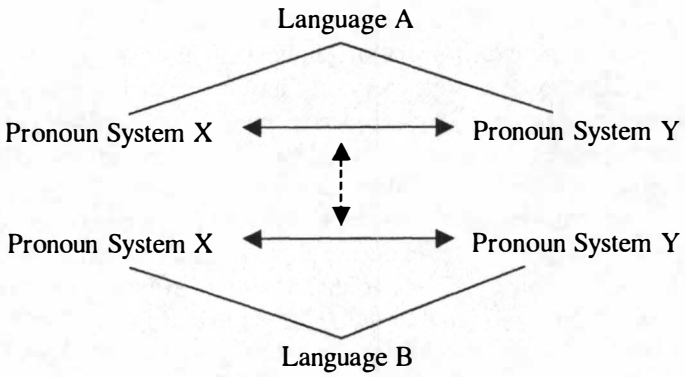



**Figure 4**: Intersystemic comparison

Once this type of investigation has been carried out for each and every one of the languages in a genetic group, it will in turn be compared across the languages. This explains why there is also a broken double-headed arrow connecting the intersystemic comparison of Language A and that of Language B in Figure 4. For instance, suppose the third person singular focus and subject pronouns in Language A and Language B have a sequence of a consonant and a vowel in common, whereas the third person singular focus and subject pronouns in Language C and Language D share a completely different sequence of a consonant and a vowel. This information may then be interpreted to be suggestive of A and B being closer to each other as opposed to C and D, or C and D being closer to each other as opposed to A and B, within the given genetic group.

Evidence has over the past decades been accumulated to the effect that the Nuclear Micronesian languages are recognised as a well-demarcated subgroup within Oceanic (Bender 1971; Bender 1984; Bender & Wang 1985; Jackson 1983, 1986; Pawley & Ross 1995), although the exact higher-level subgrouping of these languages has not yet been arrived at. Thus, it will be interesting to ascertain whether or not the personal pronoun system can be (re)scrutinised with a view to throwing some light on the internal genetic relationships of the languages, which possess as many as four different personal pronoun systems: (i) focus (also known as absolute or independent), (ii) subject, (iii) object, and (iv) possessive.[8] This is not to say, of course, that the personal pronoun system has never been utilised in Micronesian comparative linguistics for subgrouping purposes. Quite the contrary. Jackson (1983:357–363, 1986:205–207) examines the personal pronoun systems of a sizeable number of Nuclear Micronesian languages with this very goal in mind.

---

8 For the sake of convenience and comparability, I ignore here the issue as to whether the subject and object pronouns in the Nuclear Micronesian languages are referential pronouns or 'functionally ambiguous agreement markers' (Bresnan & Mchombo 1987). For detailed discussion, see Song (1994:523–547).

## 3 Internal subgrouping of Nuclear Micronesian and pronoun systems

The PMc personal pronoun system reconstructed by Jackson (1983:358–359; 1986: 205–207) is reproduced in Table 2, along with the supporting data from a number of Nuclear Micronesian languages.[9]

**Table 2:** Micronesian personal pronoun systems in Jackson (1986)

| | | *PMc | Ksr | Kir | Mrs | Pon | Mok | *PTk |
|---|---|---|---|---|---|---|---|---|
| **Focus** | | | | | | | | |
| | 1sg | **ŋau* | *nga* | *ngngai* | *ña* | *ngehi* | *ngoahi* | **ŋaŋu* |
| | 2sg | **koe* | *kom* | *ngkoe* | *kwe* | *kowe* | *koawoa* | **koe-ña* |
| | 3sg | **ia* | *el* | *ngaia* | *e* | *ih* | *ih* | **ia* |
| | 1pl.incl | **ki(t,c)a* | *kuht* | *ngaira* | *kōj* | *kit-* | *kihs* | **kica* |
| | 1pl.excl | **kamami* | *kitacl* | – | *kōmmem* | *kiht* | – | **kaamami* |
| | | **kami* | – | – | *kōm* | – | *kam-* | **kami* |
| | 2pl | **kamii* | – | *ngkamii* | *komi* | – | – | **kamii* |
| | | **kamwu* | *komtacl* | – | *kom* | *kumw-* | *kamw-* | – |
| | 3pl | **ira* | *eltahl* | *ngaiia* | *er* | *ir-, ihr* | *ihr* | **ira* |
| **Subject** | | | | | | | | |
| | 1sg | **ú* | @ | *i* | *i-* | *i* | @ | **ú* |
| | 2sg | **ko* | @ | *ko* | *ko-* | *ke* | @ | **ko* |
| | 3sg | **e* | @ | *e* | *e-* | *e* | @ | **e* |
| | 1pl.incl | **t(i,e)* | @ | *ti* | *je* | @ | @ | **Ti* |
| | 1pl.excl | **kami* | @ | – | @ | *se* | @ | **kami* |
| | 2pl | **kamwu* | @ | *kam'* | @ | @ | @ | **kamwu* |
| | 3pl | **ra* | @ | *a* | *re-* | @, *re* | @ | **re* |

9 As can be seen in Table 2, Jackson sets up doublets for the first person plural inclusive focus and object pronouns (**ki(t,c)a*), the first person plural exclusive focus and object pronouns (**kamami* and **kami*), and the second person plural focus pronoun (**kamii* and **kamwu*). He first points out that in the case of the first person plural inclusive pronoun **ki(t,c)a* the difference in the grade of the medial consonant is also reflected in other Oceanic languages (e.g. Fijian) (also see below). He (1986:205) argues, then, that the postulation of the doublets for the second person plural focus pronoun, and the first person plural exclusive focus and object pronouns is only confined to the focus pronoun system, because in PMc focus and object pronouns were not distinct in the plural, and because the PMc plural focus pronouns in fact functioned as object pronouns (Harrison 1978:1082). Jackson (1986:205) also draws attention to the fact that one member of each of the doublets is identical to the corresponding reconstructed subject pronoun. He interprets this to be suggestive of there having been a confusion in pre-PMc between focus and subject pronouns with the effect that subject pronouns were conscripted into service as focus pronouns by the time of PMc. In pre-PMc, then, **kami* and **kamwu* were subject pronouns, whereas the other members of the doublets, **kamami* and **kamii*, were focus pronouns (Jackson 1986:205).

| | *PMc | Ksr | Kir | Mrs | Pon | Mok | *PTk |
|---|---|---|---|---|---|---|---|
| **Object** | | | | | | | |
| 1sg | *-*ai* | -*yuh* | -*ai* | -*eō* | -*ie* | @ | *-*ai* |
| 2sg | *-*ko* | @ | -*ko* | -*eok* | -*uhk* | @ | *-*ko* |
| 3sg | *-*a* | @ | -*a* | -*e* | -*Ø* | @ | *-*a* |
| 1pl.incl | **ki(t,c)a* | @ | -*(i)ra* | @ | @ | @ | **kica* |
| 1pl.excl | **kamami* | @ | – | @ | @ | @ | **kamami* |
| | **kami* | @ | – | @ | @ | @ | **kami* |
| 2pl | **kamii* | @ | @ | @ | @ | @ | **kamii* |
| 3pl.HUM | **ira* | @ | -*ia* | @ | @ | @ | **ira* |
| 3pl.INA | *-*ni* | – | – | – | – | – | *-*nini* |
| 3pl.BNP | *-*xi* | – | -*i* | – | – | – | *-*i* |
| **Possessive** | | | | | | | |
| 1sg | *-*xu* | -*k* | -*u* | -*/h/* | -*i* | -*i* | *-*i* |
| 2sg | *-*mwu* | -*m* | -*m'* | -*m*, | -*mw* | -*mw* | *-*mwu* |
| 3sg | *-*ña* | -*Ø*, -*l* | -*na* | -*n* | -*Ø* | -*Ø*, -*n* | *-*ña* |
| 1pl.incl | *-*ca* | -*sr* | -*ra* | -*d* | -*t*- | -*s*- | *-*ca* |
| 1pl.excl | *-*mi* | -*ktacl* | – | -*m* | -*t* | -*m* | *-*mi* |
| | *-*mami* | – | – | – | – | – | *-*mami* |
| 2pl | *-*mii* | -*mtacl* | -*mii* | -*mi* | -*mw*- | -*mw*- | *-*mii* |
| 3pl | *-*(i)ra* | -*ltacl* | -*ia* | -*er* | -*Vr*- | -*Vr*- | *-*ira* |

**Note:** @ = the focus pronoun used for this function, BNP = before NPs, excl = exclusive, HUM = human, INA = inanimate, incl = inclusive, pl = plural, sg = singular

Jackson (1983:357–363) draws only one conclusion from the data in Table 2 for purposes of internal subgrouping, however. He points out that the Gilbertese third person plural subject pronoun reflects PEO **da* (Pawley 1972:67), whereas Proto Trukic, Ponapean, and Marshallese all reflect **re*. He then takes the form **re* to be a shared innovation (Jackson 1983:362, 435), whereby the Trukic and Ponapeic languages, and Marshallese are subsumed under PWMc, as opposed to Gilbertese and Kusaiean (see Figure 2).

The Nuclear Micronesian languages have as many as four pronoun systems, thereby exhibiting a very high degree of multiple paradigmaticity. Even if the distinction between inclusive and exclusive in the first person plural is ignored, the four pronoun systems (focus, subject, object, and possessive) will each have at least two dimensions of paradigmaticity: (i) person (first, second, and third); and (ii) number (singular, and plural).[10] But, as has been shown above, the personal pronoun systems have been little used in Nuclear Micronesian comparative linguistics. Why could this be so?

This question may perhaps best be answered by Harrison's (1978) diachronic scenario about the Micronesian personal pronoun system. He postulates that in pre-PMc the focus pronouns all functioned as object pronouns, and that the object pronouns gradually became fused with the verb to varying degrees (i.e. so-called verb–object attraction; cf. Song 1994). Thus, 'the reconstructed PMc pronoun system reflects an early stage of verb–object attraction, having begun in the singular without affecting the plural forms and moving from

10 The personal pronoun system of the Ponapeic languages has an additional number, namely dual.

third person to first person (Harrison 1978:1095).'[11] He (1978:1098) is also of the opinion that the verb–object attraction process, in full swing in PMc, went to completion 'at the peripheries of geographic Micronesia', i.e. Gilbertese, and Trukic.[12] 'Closer to the geographic core', i.e. Kusaiean, Ponapean, Marshallese, and Mokilese, on the other hand, there seem to have been 'moves in the opposite direction', i.e. some replacement of object pronouns by focus pronouns, 'along with moves in the direction of closer morpho-syntactic binding of verb and object pronoun'. He (1978:1099) suggests that what interfered with the verb–object attraction process in the core of Micronesia was the spread of final-vowel deletion, which gave rise to 'canonical shapes not amenable to the suffixation [or fusion] of object pronouns' to the verb. Said differently, in the core of Micronesia final-vowel deletion caught up with verb–object attraction, thereby not only blocking further development of verb–object attraction but perhaps also setting in motion the replacement of object pronouns by focus pronouns, whereas final-vowel deletion entered the peripheries of Micronesia (i.e. Trukic and Gilbertese) only after the verb–object attraction process had run its full course.[13]

What is intriguing about Harrison's scenario is that the process of final-vowel deletion 'pursued' that of verb–object attraction, both beginning in the geographic core of Micronesia, the eastern Carolines, and subsequently spreading outwards towards the geographic peripheries of Micronesia, and that the interaction of these two processes is claimed to have had a direct bearing on the extent of the replacement of object pronouns by focus pronouns. If this is a correct depiction of what happened in Micronesian linguistic history, then it may not come as a total surprise that the object pronoun system has not provided much insight into the internal genetic relationships of the Nuclear Micronesian languages, because the replacement of object pronouns by focus pronouns may, if anything, be regarded as more of an areal trait than a genetic one.

What about the paradigmaticity of the subject-pronoun system? As with object pronouns, 'the replacement of earlier subject pronouns by focus pronouns has occurred to various extents in all M|c| languages except for |Gilbertese| and the Trukic languages' (Jackson 1986:205). In Kusaiean and Mokilese, for example, the focus-pronoun system is used in full for subject function. As a matter of fact, such a replacement is not unheard of in the context of Oceanic languages; Ross (1988:366) points out that 'it is probable that this

---

11 Harrison's (1978:1081) reconstructed PMc focus and object pronouns, as reproduced below, are more or less similar to those reconstructed by Jackson (in Table 2) (but see Evans 1995:136–152, especially for lack of the first person plural and second person plural in POc).

| | Focus | Object |
|---|---|---|
| 1sg | **ngai* | **ai* |
| 2sg | **koe* | **ko* |
| 3sg | **ai* | **a* |
| 1pl.incl | **ki(t',t)a* | **ki(t',t)a* |
| 1pl.excl | **ka(ma)mi* | **ka(ma)mi* |
| 2pl | **kamiu* | **kamiu* |
| 3pl | **ira* | **ira* |

12 In this chapter, I assume that the Trukic languages are taken to have undergone no replacement of object pronouns by focus pronouns, as indicated in Table 2. This is not entirely correct, because it seems, for example, that in Pulo Annian the plural object pronouns have been replaced by the corresponding plural focus pronouns, whereas in Woleaian the first and second person plural object pronouns have been replaced by the corresponding focus pronouns. But it is clear from a comparison of the Trukic languages that lack of formal identity between the focus and object pronoun systems is the norm.

13 Harrison (1978:1099) thinks that some pressure to re-establish the earlier system is responsible for 'the drift back towards an absolute object pronoun system', although he is not sure as to what the source of this pressure is.

replacement process has occurred many times in the history of Central and Eastern Malayo-Polynesian languages, including those of Oceania'. But, more importantly, it seems to be confined geographically to what Harrison (1978:1098) refers to as the core area of Micronesia (e.g. the replacement being complete in Mokilese, almost complete in Kusaiean, and partial in Marshallese and Ponapean). In Gilbertese and the Trukic languages (or the geographic peripheries of Micronesia), on the other hand, there is no replacement of subject pronouns by focus pronouns in evidence. In other words, the make-up and distribution of subject pronouns in the Nuclear Micronesian languages may also not be as genetic as areal, thereby suggesting strongly that it may be injudicious to utilise the paradigmaticity of the subject pronoun system for the investigating of the internal genetic relationships of the Nuclear Micronesian languages.

The foregoing can easily be double-checked by scanning the subject and object systems across the languages in Table 2. There is a symmetry of varying degrees between the distribution of the symbol '@' (which represents the focus pronoun being used for subject or object function) in the subject pronoun system, and that in the object-pronoun system. Mokilese exhibits a complete symmetry between the two systems, Kusaiean an almost complete symmetry, and both Ponapean and Marshallese a partial symmetry.

The paradigmaticity of the subject-pronoun and object-pronoun systems, whether examined intrasystemically or intersystemically, will yield little valuable information for possible subgroupings, because the replacement of subject and object pronouns by focus pronouns in Kusaiean, Mokilese and, to a lesser extent, Marshallese and Ponapean, is more of an areal phenomenon than a genetic one.

## 4 The focus and possessive pronoun systems

The preceding discussion leaves the focus and possessive pronoun systems to be assessed for their usefulness in the understanding of the genetic relationships within Nuclear Micronesian. These systems also seem to provide little information for possible subgroupings, when studied intrasystemically, however (see Table 2).[14] From the possessive pronoun system, one may notice (i) that in the Ponapeic languages the same form *-mw(-)* is used for the second person both singular and plural; and (ii) that in Gilbertese all members of the focus-pronoun system occur with *ng-* (i.e. accretion of a velar nasal). Perhaps the first piece of information may be used in support of Ponapean and Mokilese forming the Ponapeic subgroup, which has already been well established in Micronesian comparative linguistics (e.g. Rehg & Bender 1990:24). The second point hardly bears mention as it concerns only one language. Not unexpectedly, Jackson (1983:357–363, 1986:205–207) also makes little use of these systems in his subgrouping attempt. Therefore, the focus and possessive pronoun systems seem to be as inefficacious as the other two systems.

---

14 Ponapean has second and third person singular honorific personal pronouns as well. The second person singular honorific pronouns are *komwi* (focus) and *komw* (subject) (Rehg 1981:368). These, however, bear much resemblance to the second person focus and subject pronoun *kom* in Kusaiean. It is not clear at the moment how this similarity can be explained.

However, there are three observations about the focus and possessive pronoun systems that may defy this less than positive impression. First, all possessive pronouns are suffixes, over half of them consisting of only a single consonant. This suggests strongly that the possessive pronoun system of Nuclear Micronesian is of some antiquity. It was most likely inherited from POc. Indeed, the PMc possessive pronoun system in Table 2 bears a very strong resemblance to Ross's (1988:112) POc possessive pronoun system in Table 3 (cf. Pawley 1972:61–75; Lichtenberk 1985:113).

**Table 3**: POc focus and possessive pronoun systems (Ross 1988)

| | Focus | Possessive |
|---|---|---|
| 1sg | **iau, *au* | **-gu* |
| 2sg | **iko[e], *ko[e]* | **-mu* |
| 3sg | **ia,* (?) **a* | **-ña* |
| 1pl.incl | **kita* | **-da* |
| 1pl.excl | **kami, *kai,* (?) *kamami* | **-ma[m]i* |
| 2pl | **kamu, *kau, *kamiu* | **-m[i]u* |
| 3pl | **(k)ira* | **-di[a]* |

It may thus be fair to say at least that the possessive pronoun system is much older than the subject- or object-pronoun system, which is known to have been, to varying extents, replenished with focus pronouns. Second, there is no suggestion to the best of my knowledge that the possessive pronoun system has ever been replaced by the focus or any other pronoun system (see Lichtenberk 1986:62–68; Ross 1988:208; and Evans 1995: *passim* for the opposite direction of replacement in other Oceanic languages). Finally, Jackson's PMc focus pronoun system in Table 2 bears a transparent similarity to Ross's (1988:367) reconstructed POc focus pronoun system in Table 3, thereby suggesting that there is also continuity between the focus pronoun system in POc, and those in contemporary Nuclear Micronesian languages.

The preceding observations call for an intersystemic comparison of the focus and possessive pronoun systems. To this end, two more tables are presented below. Table 4 contains the focus and possessive pronoun systems in Kusaiean (Lee 1975), Gilbertese (Groves et al. 1985), Marshallese (Bender 1969; Zewen 1977; Pagotto 1987), Ponapean (Rehg 1981), Mokilese (Harrison 1976), Trukese (Dyen 1965), Puluwat (Elbert 1974), Sonsorolese (Capell 1969), Woleaian (Sohn 1975), Pulo Annian (Oda 1977) and Ulithian (Sohn and Bender 1973) — at the risk of repeating some of the information contained in Table 2.

**Table 4:** Focus and possessive systems in Nuclear Micronesian

| | Ksr | Kir | Mrs | Pon | Mok | Trk | Pul | Sns | Wol | Pua | Uli |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Focus** | | | | | | | | | | | |
| 1 sg | *nga* | *ngngai* | *ña* | *ngehi* | *ngoahi* | *gaag* | *nga(ang)* | *ŋa:ŋ* | *gaang* | *ngangi* | *gaag* |
| 2sg | *kom* | *ngkoe* | *kwe* | *kowe* | *koawoa* | *jeen* | *yeen* | *xe̲re̲* | *geel* | *kena* | *xeel* |
| 3sg | *el* | *ngaia* | *e* | *ih* | *ih* | *jiij* | *yiiy* | *i:e̲* | *iiy* | *ia* | *yiiy* |
| 1 pl.incl | *kuht* | *ngaira* | *kōj* | *kit-* | *kihs* | *kiic* | *kiir* | *kis* | *giish* | *kisa* | *xiic, xa* |
| 1 pl.excl | *kitacl* | – | *kōmmem kōm* | *kiht* | *kam-* | *jääm* | *yææmem* | *xame̲m* | *gaamam* | *kamami* | *xaamami* |
| 2pl | *komtacl* | *ngkamii* | *komi kōm¸* | *kumw-* | *kamw-* | *jääami* | *yææmi* | *xami* | *gaami* | *kaamii* | *xaamiyi* |
| 3pl | *eltahl* | *ngaiia* | *er* | *ir-, ihr* | *ihr* | *jiir* | *yiiŕ* | *ile* | *iir* | *ila* | *yiir* |
| **Possessive** | | | | | | | | | | | |
| 1 sg | *-k* | *-u* | *-/h/* | *-i* | *-i* | *-i, -j* | *-y(i)* | *-i* | *-i* | *-i* | *-yi* |
| 2sg | *-m* | *-m'* | *-m¸* | *-mw* | *-mw* | *-(V)b* | *-mw* | *-m* | *-mw* | *-mwu* | *-mu* |
| 3sg | *-Ø, -l* | *-na* | *-n* | *-Ø* | *-Ø, -n* | *-n* | *-n, -y* | *-r$^a$* | *-l* | *-na* | *-la* |
| 1 pl.incl | *-sr* | *-ra* | *-d* | *-t-* | *-s-* | *-c* | *-r* | *-s* | *-sh* | *-sa* | *-ca* |
| 1 pl.excl | *-ktacl* | – | *-m* | *-t* | *-m* | *-m* | *-mem -mám* | *-memi* | *-mam* | *-mami* | *-mami* |
| 2pl | *-mtacl* | *-mii* | *-mi* | *-mw-* | *-mw-* | *-mi* | *-mi* | *-mi* | *-mi* | *-mii* | *-miyi* |
| 3pl | *-ltacl* | *-ia* | *-er* | *-Vr-* | *-Vr-* | *-r* | *-ŕ* | *-l* | *-r, -l* | *-ila* | *-yire* |

Table 5 provides a schematic summary of the intersystemic comparison of the focus and possessive pronoun systems in each of these Nuclear Micronesian languages.[15]

**Table 5:** Intersystemic comparison of focus and possessive pronoun systems

| | | |
|---|---|---|
| KUSAIEAN: | | <[F] SG ⊇ [P] SG, except [F] 1SG ≠ [P] 1SG> & <[F] PL ⊇ [P] PL, except [F] 1PL.INC ≠ [P] 1PL.INC> |
| MARSHALLESE: | | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL, except [F] 1PL.INC ≠ [P] 1PL.INC> |
| GILBERTESE: | | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL, but no distinction between INC and EXC in 1PL> |
| PONAPEIC: | | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| TRUKIC: | Trukese: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| | Puluwat: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| | Sonsorolese: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| | Woleaian: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| | Pulo Annian: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |
| | Ulithian: | <[F] SG ≠ [P] SG> & <[F] PL ⊇ [P] PL> |

**Note:** 'X ⊇ Y' means that Y is contained within X to varying degrees ranging from partial formal similarity to complete formal identity; 'X ≠ Y' means 'X is dissimilar formally to Y'; '[F]' focus pronoun system; '[P]' possessive pronoun system; EXC = exclusive; INC = inclusive; SG = singular; PL = plural.

There are three points emerging from Table 5 that merit discussion. First, in Gilbertese, and also the Ponapeic and Trukic languages, the paradigmatic members of the plural focus pronoun system have something in common with the corresponding paradigmatic members of the plural possessive pronoun system. In Mokilese, for example, the plural focus pronouns 'contain' the plural possessive pronouns on a one-to-one basis, i.e. focus 1PL.INC *kihs* ~ possessive 1PL.INC *-s-*; focus 1PL.EXC *kam-* ~ possessive 1PL.EXC *-m-*; focus 2PL *kamw-* ~ possessive 2PL *-mw-*; focus 3PL *ihr* ~ possessive 3PL *-r-*. Although Gilbertese differs from the other Nuclear Micronesian languages in that it lacks the inclusive–exclusive distinction in the first person plural throughout the personal pronoun systems,[16] it does also maintain the property of <[F] PL ⊇ [P] PL>.[17] This particular property, however, is not fully exhibited by Kusaiean and Marshallese, because in these two languages there is lack of a formal similarity between the first person plural inclusive focus and possessive pronouns (i.e. <[F] 1PL.INC ≠ [P] 1PL.INC>), although the remainder of the focus and possessive pronoun systems do display a formal similarity. This raises the question as to which of the two, <[F] 1PL.INC ⊇

15 The third person singular possessive pronoun *-n* in Puluwat is in free variation with *-y* (Elbert 1974:35). Thus, it may be said that Puluwat is different from the other Trukic languages in that there is a formal similarity between the pronoun in question and the corresponding third person singular focus pronoun, *yiiy*. But Elbert (1974:35) points out that *-n* is more common in citation forms. I will thus take *-n*, not *-y*, to be the basic form for the third person singular possessive pronoun in Puluwat.

16 Harrison (1978:1100–1101) points out that the inclusive, not exclusive, forms have been retained in the Gilbertese first person plural. This also seems to be the position of Jackson (1983, 1986).

17 Also note that most of the focus pronouns in Gilbertese, minus the initial *ng-*, can without difficulty be related to the corresponding forms in the Ponapeic and Trukic languages via regular sound correspondences (Jackson 1983:202–203).

[P] 1PL.INC> or <[F] 1PL.INC ≠ [P] 1PL.INC>, is an innovation (or a retention). In Jackson's PMc focus pronoun system in Table 2, there is a variation between **kita* and **kica* in the first person plural inclusive focus pronoun (i.e. **ki(t,c)a*). Thus, the reconstructed PMc first person plural inclusive possessive pronoun **-ca* may or may not bear resemblance to the corresponding PMc focus pronoun, depending upon which of the two, **kita* or **kica*, is taken as the PMc form. In Ross's (1988:367) reconstructed POc focus and possessive pronoun systems in Table 3, however, there is no such variation, i.e. **kita*; there is lack of a formal similarity between the first person plural inclusive focus and possessive pronouns (i.e. **kita* vs **-da*). If Ross's POc reconstruction is correct, then there is a strong possibility that the PMc form should only be **kita*, not both **kita* and **kica*. This in turn suggests that between PMc and present-day Nuclear Micronesian there may have occurred a change from <[F] 1PL.INC ≠ [P] 1PL.INC> to <[F] 1PL.INC ⊇ [P] 1PL.INC> in some Nuclear Micronesian languages, but not in others. Thus, the property of <[F] 1PL.INC ⊇ [P] 1PL.INC> may well be a post-PMc innovation shared by Gilbertese, and the Ponapeic and Trukic languages, whereas Kusaiean and Marshallese may have retained the POc property of <[F] 1PL.INC ≠ [P] 1PL.INC>. (Indeed, Kusaiean *kuht* ([F] 1PL.INC) and *-sr* ([P] 1PL.INC), and Marshallese *kōj* ([F] 1PL.INC) and *-d* ([P] 1PL.INC) seem to have descended from the respective POc forms (see Table 3) via regular sound correspondences (Jackson 1986:202–203).) This innovation, however, does not fit in comfortably with Jackson's tree model, wherein Marshallese is interposed between Gilbertese on the one hand, and Ponapeic and Trukic on the other. In view of the innovation, Gilbertese, Ponapeic and Trukic are expected to be much closer to one another than they are represented in the tree model.

The difference between Jackson's (1986) PMc **kita* and **kica* actually is the alternation between the oral (PMc **t* < POc **t*) and nasal (PMc **c* < POc **nt*) grade, which is a long-standing issue in Oceanic linguistics (see Grace 1959, 1990; Biggs 1965; Lynch 1975 and Geraghty 1983 and Ross 1988 *inter alia*). The oral grade is reflected in Kusaiean and Marshallese, whereas the nasal grade is manifested in the other Nuclear Micronesian languages. This variation in grade also happens to be witnessed elsewhere in Oceanic (Jackson 1986:205). Thus, one may argue that not much subgrouping significance can be imputed to it. In fact, if Jackson's PMc alternation between **kita* and **kica* is correct, it may be possible to say, contrary to Ross (1988:367), that the same alternation may have been present also in the POc first person plural inclusive focus pronoun (cf. Grace 1990). However, the oral/nasal alternation evident in the first person plural inclusive focus pronoun in present-day Nuclear Micronesian may well be a secondary development, thereby not reflecting a continuation of the POc alternation at all. There is some evidence in favour of this view. First, the environment in which the oral-to-nasal change in Gilbertese, Ponapeic and Trukic has occurred is very limited. This suggests that it is within the realm of possibilities that in Gilbertese, Ponapeic and Trukic the first person plural inclusive focus pronoun may have substituted the nasal grade for the oral one by analogy with the first person plural inclusive possessive pronoun. This scenario of analogical levelling also seems to be well motivated in view of the fact that in Oceanic replacement of non-possessive pronouns by possessive pronouns is known to be relatively common (e.g. Ross 1988:208, 277–278; Evans 1995). The innovation shared by Gilbertese, Ponapeic and Trukic, as opposed to Kusaiean and Marshallese, can thus be characterised by the 'spreading' of the nasal grade from the first person plural inclusive possessive pronoun to the corresponding focus pronoun. Similar secondary changes have been attested in other Oceanic languages. For instance, Ross (1988:35) is of the view that acquisition of the nasal grade reflex *g-*, rather than the expected **k* or zero, in the three disjunctive pronouns in Tabar, Notsi and

Lihir (all central New Ireland languages) 'involves some kind of analogical levelling in the pronoun paradigm, and has nothing to do with POc'. The basis of this conclusion is also the limited environment of the change from the oral to the nasal grade. Moreover, Lynch (1975:87–88) discusses a very different secondary source of the oral/nasal alternation in other Oceanic languages, namely the fusion of a preposed article consisting of a nasal consonant and a vowel with the following verb. Also see Geraghty (1983:72–96) for a phonetically motivated secondary development of the oral/nasal alternation in Eastern Fijian. Thus, the post-PMc change from the oral to the nasal grade in Gilbertese, Ponapeic and Trukic does not seem to be implausible.

Second, Kusaiean stands out from the rest, because in this language the second or third person singular possessive pronoun is identical to the final consonant of the second or third person singular focus pronoun respectively. This may be a post-PMc innovation confined to Kusaiean. The first person singular possessive pronoun in the Ponapeic languages, *-i*, may perhaps also be related formally to the corresponding first person singular focus pronoun. But I am more inclined to think that it is a reflex of the PMc first person singular possessive pronoun **-xu*, as in the case of Trukic. Thus, it may well have derived from the vowel of PMc **-xu*, rather than being identical formally to the final vowel of the first person singular focus pronoun (*ngehi* and *ngoahi* in Ponapean and Mokilese, respectively).

Finally, the lack of the first person plural inclusive–exclusive distinction in Gilbertese also seems to be a post-PMc innovation. This and the property of <[F] non-1SG ⊇ [P] non-1SG> in Kusaiean, however, contribute little to internal subgrouping, because they are shared by no other language(s).

## 5 Conclusion

I have attempted to address two important points which have recently been raised in the literature on (Micronesian) comparative linguistics. The first is Rehg's (1995) observation that adequacy of tree models for an understanding of genetic relationships can perhaps be determined on the basis of the careful tracking of the distribution of all innovations 'without regard to preconceived notions of language and subgrouping boundaries'. The second is Nichols's (1996) demonstration of the role of paradigmaticity as probative evidence for genetic relatedness. Nichols's (and Meillet's) scepticism of personal pronouns being such evidence notwithstanding, I have suggested that phonosymbolism of personal pronouns can be 'checked' if and when comparative work is carried out on a given group of languages which are already known — on the basis of other probative evidence — to have emerged from a common source (although their internal relationships may be far from established). With these points in mind, I have carried out an intersystemic examination of the focus and possessive personal pronoun systems of the Nuclear Micronesian languages. The conclusion turns out to be somewhat at odds with Jackson's (1983, 1986) stratified tree model of Nuclear Micronesian, because, although it forms PCMc with Gilbertese, Ponapeic and Trukic, Marshallese does not share the innovation, namely the formal similarity between the first person plural inclusive focus and possessive pronouns, with those languages.

## References

Bender, Byron W., 1969, *Spoken Marshallese.* Honolulu: University of Hawaii Press.

—— 1971, Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol. 8: *Linguistics in Oceania*, 426–465. The Hague: Mouton.

—— 1984, Object marking in Marshallese. In Byron W. Bender, ed. *Studies in Micronesian linguistics*, 443–465. Canberra: Pacific Linguistics.

Bender, Byron W. and Judith W. Wang, 1985, The status of Proto-Micronesian. In Pawley and Carrington, eds 1985:53–92.

Biggs, Bruce, 1965, Direct and indirect inheritance in Rotuman. *Lingua* 14:383–415.

Blake, Barry, 1989, Redefining Pama-Nyungan: towards the prehistory of Australian languages. *Aboriginal Linguistics* 1:1–90.

—— 1990, The significance of pronouns in the history of Australian languages. In Philip Baldi, ed. *Linguistic change and reconstruction methodology*, 435–450. Berlin: Mouton de Gruyter.

Bloomfield, Leonard, 1933, *Language.* New York: Henry Holt.

Bresnan, Joan and Sam A. Mchombo, 1987, Topic, pronoun, and agreement in Chicheŵa. *Language* 63:741–782.

Capell, Arthur, 1969, *Grammar and vocabulary of the language of Sonsorol-Tobi.* Sydney: University of Sydney.

Dyen, Isidore, 1965, *A sketch of Trukese grammar.* New Haven: American Oriental Society.

Elbert, Samuel, H., 1974, *Puluwat grammar.* Canberra: Pacific Linguistics.

Evans, Bethwyn, 1995, Reconstructing object markers in Oceanic languages. BA Honours sub-thesis, The Australian National University.

Geraghty, Paul A., 1983, *The history of the Fijian languages.* Honolulu: University of Hawaii Press.

Geraghty, Paul, Lois Carrington and S.A. Wurm, eds, 1986, *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics.* Canberra: Pacific Linguistics.

Grace, George W., 1959, *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family.* Memoir 16, International Journal of American Linguistics.

—— 1986, Further thoughts on Oceanic subgrouping. In Geraghty, Carrington and Wurm, eds 1986:1–12.

—— 1990, 'Consonant grade' in Oceanic languages. In J.H.C.S. Davidson, ed. *Pacific island languages: essays in honour of G.B. Milner*, 41–49. London: School of Oriental and African Studies, University of London.

Groves, Terab'ata R., Gordon W. Groves, and Roderick Jacobs, 1985, *Kiribatese: an outline description.* Canberra: Pacific Linguistics.

Harrison, Sheldon P., 1976, *Mokilese reference grammar.* Hawaii: University Press of Hawaii.

—— 1978, Transitive marking in Micronesian languages. In Stephen A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 1067–1127. Canberra: Pacific Linguistics.

Irwin, Geoffrey, 1992, *The prehistoric exploration and colonization of the Pacific.* Cambridge: Cambridge University Press.

Jackson, Frederick H., 1983, The internal and external relationships of the Trukic languages of Micronesia. PhD dissertation, University of Hawai'i.

—— 1986, On determining the external relationships of the Micronesian languages. In Geraghty, Carrington and Wurm, eds 1986:201–238.

Lee, Kee-dong, 1975, *Kusaiean reference grammar*. Honolulu: University Press of Hawaii.

Lichtenberk, Frantisek, 1985, Possessive constructions in Oceanic languages and in Proto-Oceanic. In Pawley and Carrington, eds 1985:93–140.

—— 1986, Syntactic-category change in Oceanic languages. *Oceanic Linguistics* 24:1–84.

Lynch, John, 1975, Oral/nasal alternation and the realis/irrealis distinction in Oceanic languages. *Oceanic Linguistics* 14:87–99.

Meillet, Antoine, 1958, *Linguistique historique et linguistique générale*. Société Linguistique de Paris, Collection Linguistique 8. Paris: Librairie Honoré Champion.

Nichols, Johanna, 1996, The comparative method as heuristic. In Mark Durie and Malcolm Ross, eds *The comparative method reviewed: regularity and irregularity in language change*, 39–71. New York: Oxford University Press.

Oda, Sachiko, 1977, The syntax of Pulo Annian: a Nuclear Micronesian language. PhD dissertation, University of Hawai'i.

Pagotto, Louise, 1987, Verb subcategorization and verb derivation in Marshallese: a localistic lexicase analysis. PhD dissertation, University of Hawai'i.

Pawley, Andrew, 1972, On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Studies in Oceanic culture history*, vol. 3, 1–142. Honolulu: Bernice P. Bishop Museum.

Pawley, Andrew and Lois Carrington, eds, 1985, *Austronesian linguistics at the 15th Pacific Science Congress*. Canberra: Pacific Linguistics.

Pawley, Andrew and Malcolm Ross, 1995, The prehistory of the Oceanic languages: a current view. In Peter Bellwood, James J. Fox and Darrell Tryon, eds *The Austronesians*, 39–74. Canberra: Department of Anthropology, The Australian National University.

Rehg, Kenneth L., 1981, *Ponapean reference grammar*. Honolulu: University Press of Hawaii.

—— 1995, The significance of linguistic interaction spheres in reconstructing Micronesian prehistory. *Oceanic Linguistics* 34:305–324.

Rehg, Kenneth L. and Byron W. Bender, 1990, Lexical transfer from Marshallese to Mokilese: a case of intra-Micronesian borrowing. *Oceanic Linguistics* 29:1–26.

Ross, M.D., 1988, *Proto Oceanic and the Austronesian languages of Western Melanesia*. Canberra: Pacific Linguistics.

—— 1996, Contact-induced change and the comparative method: cases from Papua New Guinea. In Mark Durie and Malcolm Ross, eds *The comparative method reviewed: regularity and irregularity in language change*, 180–217. New York: Oxford University Press.

Sohn, Ho-min, 1975, *Woleaian reference grammar*. Honolulu: University Press of Hawaii.

Sohn, Ho-min and Byron W. Bender, 1973, *A Ulithian grammar*. Canberra: Pacific Linguistics.

Song, Jae Jung, 1994, The Verb-Object Bonding Principle and the pronominal system: with special reference to Nuclear Micronesian languages. *Oceanic Linguistics* 33:517–565.

Southworth, Franklin, C., 1964, Family-tree diagrams. *Language* 40:557–565.

Zewen, François X.N., 1977, *The Marshallese language: a study of its phonology, morphology and syntax*. Berlin: Dietrich Reimer.

Song, J.J. "Internal subgrouping and pronominal paradigmaticity: the case of Nuclear Micronesian". In Bowden, J. and Himmelmann, N. editors, *Papers in Austronesian subgrouping and dialectology*.
**PL-563:111-128**. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.111*

# 7 *Notes on the southern Muna dialect*

RENÉ VAN DEN BERG

## 1 Introduction

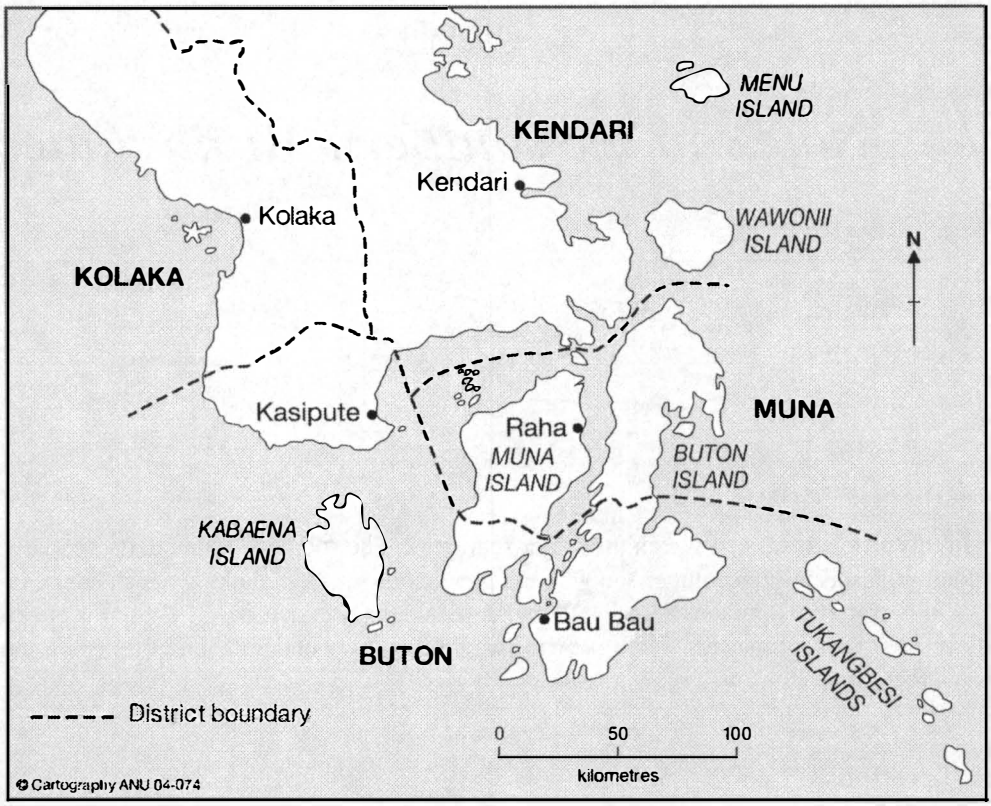
In this article I want to sketch the major features of the southern Muna dialect (Sulawesi, Indonesia), especially in comparison with the prestigious northern dialect, which has been the basis of a grammar (van den Berg 1989) and a dictionary (van den Berg 1996). Before going into its structural properties, I will present the geographical and sociolinguistic situation of the southern dialect.

### 1.1 Location

The Muna language is spoken on the whole island of Muna, off the southeast coast of Sulawesi, Indonesia. Muna speakers are also found on the west coast of Buton and on the islands of Kadatua and Siompu, southwest of Baubau, the capital of Buton. There are also sizeable communities of Muna speakers in Kendari and — until recently — in Ambon. The total number of speakers of the language can be estimated at around 250,000 and possibly up to 300,000.

The main dialect, which I have called Standard Muna, is spoken in the northern two-thirds of the island of Muna, while the southern Muna dialect, dealt with here, is spoken in the remaining part of the island. The dialect boundary separating these two varieties actually closely follows an administrative boundary, namely the *kabupaten* (regency or district) boundary. Somewhat confusingly, *kabupaten* Muna consists of the northern two-thirds of the island of Muna and the northern half of the island of Buton, while the southern one-third of the island of Muna and the southern half of Buton island belong to *kabupaten* Buton. Other areas which belong to *kabupaten* Buton are the Tukangbesi islands (frequently designated by their Indonesian acronym Wakatobi), the island of Kabaena west of Muna and a section of the mainland of southeast Sulawesi, viz. the *kecamatan* (subdistricts) Poleang and Rumbia. (See Map 1: Southeast Sulawesi.) For this reason, a southern Muna speaker will often call himself Butonese (*orang Buton*), since administratively the area he or she is from is part of *kabupaten* Buton, a prestigious place which is known all over Indonesia. In contrast, few people outside of Southeast Sulawesi have ever heard of Muna.

John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 129–170.
Canberra: Pacific Linguistics, 2004.

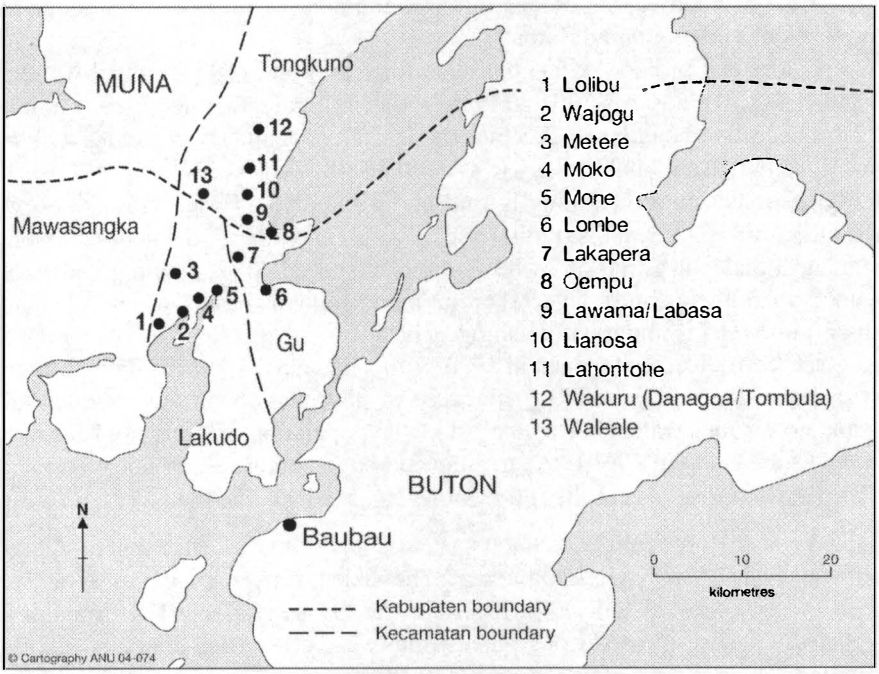



**Map 1:** Southeast Sulawesi

Before 1990 the southern one third of Muna (part of *kabupaten* Buton) was made up of two subdistricts: Gu and Mawasangka, collectively called Gumas. Currently there are three subdistricts in this area: Gu, Lakudo and Mawasangka, sometimes referred to as Gulamas. This is the area where the southern dialect is spoken. This southern dialect area is actually not uniform, although a detailed comparision of the subdialects remains to be worked out. However, the difference between northern and southern Muna is very obvious, marked by a bundle of three important phonetic isoglosses and a cognate percentage of 87% (see below for details). There are even specific terms for speaking the northern dialect, which is referred to by the verb *defokaru* and speaking the southern dialect, termed *defo'ae*. Intelligibility between the dialects is fairly low, especially on first contact.

The southern dialect under discussion here may be called the Gu subdialect of the southern Muna dialect, since the basis of the description is the speech variety in the village of Lakapera in Gu. But again the dialect boundary and subdistrict boundary do not coincide [see Map 2: Southern Muna]. This Gu subdialect is spoken in *kecamatan* Gu in the villages of Lakapera and the township of Lombe, which consists of the administrative villages Bombonawulu, Watulea and Walando. In *kecamatan* Lakudo the southern dialect described here is spoken in the villages of Lolibu, Mone, Moko and Wajogu. It is also spoken in the southernmost villages of *kecamatan* Tongkuno, viz. the village of Waleale (recently divided into the villages of Waleale, Kulidawa and Matombura), the southern part of Lawama (recently divided into the villages of Labasa and Lawama), and part of Oempu (usually referred to by its traditional name Walingkabola), all located north of the administrative

boundary in *kabupaten* Muna. Pockets of southern speakers are also found around Wakuru, the capital of *kecamatan* Tongkuno. In other words, the dialect boundary runs almost parallel to the administrative *kabupaten* boundary, with the big exception of Waleale, parts of Lawama and Oempu, villages just north of that boundary. I will refer to these areas as the 'enclaves'. The reverse situation is true for the village of Metere, firmly located in *kecamatan* Lakudo, but speaking the northern dialect.




**Map 2:** Southern Muna

My data on other parts of the southern dialect are limited; *kecamatan* Mawasangka has its own subdialect, while the rest of *kecamatan* Lakudo has some unique features, as do the dialects of Siompu and Kadatua. The situation on the west coast of Buton deserves further study.

## 1.2 History

Couvreur (1935, 2001) contains a brief account of the history of Muna. According to oral information obtained by Couvreur, which is still recounted today with some variations, the original population of Muna descended from the crew of one of Sawerigading's ships which ran aground on the coral island. They discovered a stone that had flowers growing on it (*kontu kowuna*), which gave the island its name *Wuna*, 'flower', of which the exonym Muna is a corruption. Oral tradition has it that the first inhabitated places on Muna were the villages of Wamilei and Tongkuno in the central-eastern hill country, after which the population gradually moved west, south and north. Muna also acquired a royal dynasty, of

which the progenitor emerged from a bamboo stalk, according to legend. One of the later kings, Lakilaponto, organised Muna into 36 different political units, roughly corresponding to today's villages. He also helped the neighbouring island of Buton by killing the fearful pirate La Bolontio, after which he become king of Buton, known by the name Murhum. During his reign he embraced Islam, which was introduced from Ternate, and Murhum was the first to call himself Sultan of Buton. When he became king of Buton, he gave up power over Muna, but took with him the two villages Bombonawulu and Lakudu (Couvreur 1935:7), in present-day *kecamatan* Gu and Lakudo. Since that time these areas no longer form part of the Muna area, but belong to the Buton area. That is why the southern third of Muna now belongs to *kabupaten* Buton.

Couvreur does not provide dates, but according to Vonk (1937:165), Murhum reigned from 1538–1584. In any case this story provides an interesting historical tradition on the current administrative boundary. The fact that a dialect boundary coincides with a political boundary is well known in dialectology (e.g. Petyt 1980:62–67).

A few questions remain. Nothing is said about the area of Mawasangka — was it not inhabited at that time? Also, this tradition seems to run counter to the general axiom that the area of greatest dialectal variation is the centre of dispersal. This would point to the area of Gu-Lakudo, where the greatest dialectal variation on Muna is found, whereas the central and northern part of Muna is remarkably homogeneous.

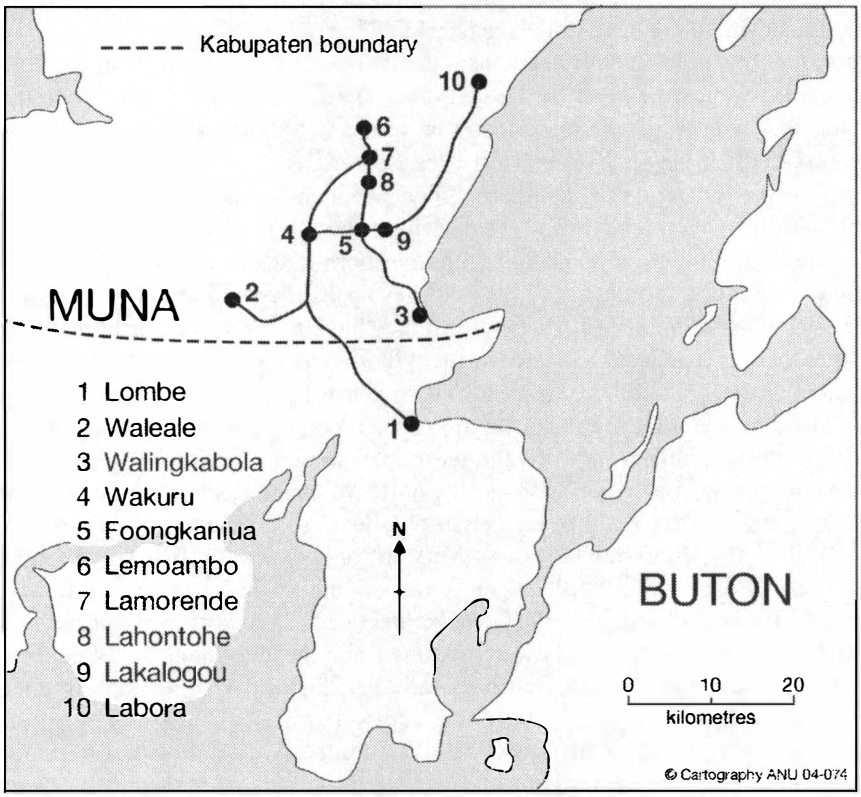
In the past, several other villages in *kecamatan* Tongkuno (north of the administrative boundary) spoke the southern dialect. Because of poor soil conditions, frequent attacks of malaria and continuous water shortage, most of the population of Tongkuno has been moved to more fertile areas in the 1960s. As mentioned above, almost all of Tongkuno was part of the northern dialect area, except the following villages (which like the others, no longer exist):

- Labora (locally pronounced as /labo:xa/), a village some 17 km north of Wakuru in *kecamatan* Tongkuno, close to the sea. The dialect spoken there may actually have been the Lakudo subdialect, rather than the Gu subdialect. These people have reportedly ended up in various places; many have resettled south of Kendari in Tanjung Tiram (*kecamatan* Moramu); others have gone to Tampunabale and other places on the west coast of Buton such as Pure and Maligano; some have gone to Oempu.
- Old Lahontohe (actual pronunciation /laghontoghe/). Reportedly many people spoke the southern dialect; they have been relocated to present-day Wakuru (the villages of Danagoa and Tombula).
- Foongkaniua, some 2 km towards Lemoambo and Oempu; people relocated to the modern villages of Lahontohe and Tombula (Wakuru).

Other villages which have been relocated from this general area have always been speaking the northern dialect, such as Lamorende, Lakologou and Lemoambo. It should also be noted that during the colonial period Waleale belonged to the subdistrict of Kabawo.

The exact distribution of the dialect situation before 1960 is still unclear and will require substantial local research. My tentative conclusions are shown on Map 3, which was drawn up with the help of government officials in Wakuru. One of the problems in trying to reconstruct the situation is that there are few if any good maps of the area. Many local names do not appear on any map. An old Dutch map of 1917 is useful, but it has obvious errors.

A second problem is the confusion over old and new village names and the people they represent. To take just one example, the old village of Lemoambo was deserted in 1967; its former location is indicated on the 1917 map. The area is currently deserted. The people from this village have scattered to at least three locations; some have been relocated in the village of Lianosaa (south of Wakuru), others have gone to Bonea and Labunti, north of Raha, reportedly fleeing the malaria-infested area they came from, and some have settled near Guali, in the northwestern corner of Muna. In all four places the old village name is in official use again (either as a village name or as a neighbourhood name), but many of the people in these places are from other areas as well.




**Map 3:** Southern Muna (pre-1960)

Another question which remains to be answered is to what extent the southern dialect spoken in these northern villages was identical to the one described here. All informants agree on the phonological features to be discussed below, but the lexical variation may actually have been much less. For example, the northern word for 'salt' is *ghohia*, in the south it is *gaha*, but reportedly southern dialect speakers in places like Lahontohe said *'o'ia*, using southern phonology on a northern lexical item.

## 1.3 Sociolinguistic situation: language status and language use

The northern dialect has always been more prestigious than the southern one. This is (as usual) a reflection of the political situation; the capital of the Muna kingdom used to be Kota Muna in central Tongkuno, right in the northern dialect area. This is where the king and the nobility lived, probably from somewhere in the 16th century until the end of the 19th century. This fortified town was deserted early in the 20th century after an internal power struggle (Couvreur 1935:11–14). The Dutch colonial government made Raha on the northeast coast the capital, but for many people Tongkuno remains an area associated with past glory and grandeur. As mentioned before, in the 1960s almost the whole population of northern and central Tongkuno was relocated to other parts of Muna. Consequently the traditional heartland of Muna has now long been empty, although some people still venture out to farm these arid areas, and there are even plans to rebuild the old kota Muna.

Reportedly, the nobility always spoke the northern dialect, even in villages such as Lahontohe where many or most of the common population spoke the southern dialect. It is unclear to what degree people understood or spoke two dialects in these enclave villages, although reportedly it may have been fairly widespread.

Because of the relocations of southern dialect speakers from Lahontohe and Foongkaniua some 30 years ago to their current areas, where they are surrounded by northern dialect speakers, children now grow up speaking the northern dialect. Several adults I interviewed who were born in old Lahontohe told me that they spoke the southern dialect when they grew up; since relocating (somewhat further south!) they have adopted the northern dialect. It thus appears that there is a dialect shift under way in these relocated areas. Careful observation of actual dialect usage would be necessary to fully understand the mechanism of this shift.

It was also reported that in the border area speakers frequently adapt their dialect to their conversation partner. From my own limited observations, I have mostly seen examples of southern Muna speakers attempting to use the northern dialect. Some interviewed people also claimed newcomers to the Wakuru area attempt to learn some southern features.

Almost all of the Muna population is Muslim; only 1–2% profess to be Christians and they are found in the southern villages of Waleale/Kulidawa/Matombura, Labasa/Lawama, Lakapera and Lolibu, all speaking the Gu subdialect. Mission work was started in the 1930s in central Muna, but lasting results were obtained only in the village of Waleale and in the capital Raha. From Waleale catholicism spread to Lolibu in the 1950s, a poor village on the southern coast with little arable soil. In the 1960s and early 1970s a Belgian priest led a local transmigration project, relocating dozens of families from Lolibu to a more fertile area, the present-day villages of Lakapera and the southern part of Labasa/Lawama. It is the speech variety of this group that is the basis for the following observations.

Knowledge of Indonesian is widespread and growing all over Muna. The resulting bilingualism is not endangering Muna, although one does meet families where the parents speak to their children in Indonesian in an attempt to help them get ready for school. From my own observations it is fairly rare, however, to meet teenage children who do not speak and understand Muna.

An important question for the future dialect situation is the role of the schools. Since about 1995 the government has promoted the use of selected regional languages in primary and junior secondary schools under the heading of *muatan lokal* ('local material'). In all of *kabupaten* Muna this is the Muna language in its standard form, the northern dialect. This means that in the southern 'enclaves' of Waleale, Labasa and Oempu children are learning to read and write the northern dialect. To what extent this will influence spoken language

remains an open question. Time devoted to *muatan lokal* is rather limited, but the exposure to a written norm will almost certainly have some influence. Since *kecamatan* Gu, Lakudo and Mawasangka form part of *kabupaten* Buton, the language taught for *muatan lokal* in these places is Wolio, the prestigious language of the old Buton sultanate (Anceaux 1952, 1987), which has been chosen for *muatan lokal* in *kabupaten* Buton. However, very few people in *kecamatan* Gu have a speaking knowledge of Wolio, and the effect of this language policy is still unknown.

Vernacular literature in the southern dialect is very limited. Whereas several books have been published in the northern dialect, of both a scholarly and a popular nature, efforts in the south have been (as far as I know) the result of one author, Lukas Atakasi, a school teacher from Lakapera. He has written the animal book *Kadadi'i ne witeno Wuna* — also published in a northern dialect version — and a retelling of the Joseph story from the book of Genesis, *Tula-tulano Yusuf*. He also adapted a short trilingual conversation book (Muna–Indonesian–English) into the southern dialect and wrote a transition primer *Poguhu kaita daebuhi wamba Wuna* ('Let's learn to write Muna') which is used in short courses for people who are literate in Indonesian and want to learn to read and write the southern Muna dialect. Although circulation of these books is limited (and largely restricted to Christian circles), they have helped to create a standardised spelling system for the southern dialect and an increased awareness of the difference between northern and southern Muna.

Currently a Bible translation project is underway in the southern dialect, with a few books already available in trial editions. These publications have further enhanced the status of the dialect.

In summary, it appears that the northern dialect is thriving, with adequate scholarly documentation, a growing body of literature, and official recognition as a school subject. The southern dialect, on the other hand, seems to be losing some ground in the northern 'enclaves', has not yet been documented, is not officially recognised, with the main efforts at standardisation and literature production the work of individuals and the church.

After this introductory section we will now look at the structural features of the southern dialect (phonology, morphology and syntax), focusing on those aspects that differ from Standard Muna, the northern dialect.

## 2 Phonology

Since the most conspicuous differences between north and south Muna are in the area of consonantal phonology, I will treat this topic in some detail. The dialects are in complete agreement on the other phonological parameters, namely:

- five vowels /i, e, a, o, u/;
- penultimate stress;
- syllable structure only V and CV (with prenasalised consonants interpreted as units).

### 2.1 Consonants

Table 1 displays the southern Muna consonants.

**Table 1:** Southern Muna consonant phonemes

| | bilabial | labio-dental | dental | alveolar | velar | glottal |
|---|---|---|---|---|---|---|
| stop voiceless | *p* | | *t̪* | | *k* | *ʔ* |
| stop voiced | *b* | | *d̪* | | *g* | |
| stop voiceless prenasalised | *mp* | | *nt̪* | | *ŋk* | |
| stop voiced prenasalised | *mb* | | | *nd* | *ŋg* | |
| implosive | *ɓ* | | | *ɗ* | | |
| nasal | *m* | | | *n* | *ŋ* | |
| fricative | | *f* | | *s* | | *h* |
| fricative prenasalised | | | | *ns* | | |
| trill | | | | *r* | | |
| lateral | | | | *l* | | |
| approximant | *w* | | | | | |

## 2.2 Allophonic variation

The dental phonemes /t̪/ and /d̪/ are realised as regular laminodental stops before the vowels /a, o, u/. They each show important allophonic variation before the high vowels /i/ and /u/.

- /t̪/ is realised as a postalveolar stop with slightly affricated release; the release point of articulation varies from alveolar via postalveolar to prepalatal. In addition, the release can be accompanied by light aspiration:

  Allophonic variants: [tʃ, tʃʰ, tˢ, tˢʰ, c, cʰ].

  Examples:

  | | |
  |---|---|
  | [fitʃu] | 'seven' |
  | [tʃʰu:] | 'knee' |
  | [tˢitˢi] | 'breast' |
  | [ɗatʃi] | 'teak' |

- /d̪/ shows similar allophonic variation before the high vowels /i/ and /u/. In this case the phonetic realisation is a postalveolar or prepalatal affricate, not unlike the Indonesian phoneme <j>, which is usually regarded as a patalal plosive. In the Muna case, I would prefer to call this allophone a laminal postalveolar affricate [dʒ].

  Examples:

  | | |
  |---|---|
  | [ɓadʒu] | 'shirt' |
  | [dʒudʒu] | 'push' |
  | [dʒini] | 'evil spirit' |
  | [ɓihidʒiŋi] | 'mock' |

The allophonic nature of [dʒ] was not recognised in van den Berg (1991a:33), where I assumed [dʒ] (there written as /j/) was a separate phoneme in the Lombe word list, although it was found only before /u/.

A comparison with the chart for the northern dialect as published in van den Berg (1989:16) shows some other differences, which are mostly of a minor nature:

- /ɗ/ has a stronger implosive quality in the south than in the north; hence it is subsumed under implosives, rather than under voiced stops;
- the northern dialect has the palatal stops /c, ɟ/ and the palatal approximant /y/ occurring in a limited number of loan words; in the south such words are even less frequent and can best be considered as unadapted loans from Indonesian. Examples are /cama/ 'subdistrict head' and /yakini/ 'conviction', from Indonesian *camat* and *yakin* respectively.

## 2.3 Orthographic conventions

In this section I will introduce the southern Muna orthography as it is currently in use, which will then be employed throughout the rest of this paper. Most phonemes are orthographically represented in the traditional way, and for the remaining phonemes the following conventions are used:

| **phoneme** | **grapheme** |
|---|---|
| /t̪/ | *t* |
| /d̪/ | *dh* |
| /ɓ/ | *bh* |
| /ɗ/ | *d* |
| /ʔ/ | ' |
| /n̪t̪/ | *nt* |
| /ŋ/ | *ng* |
| /ŋk/ | *ngk* |
| /ŋg/ | *ngg* |

Notice that for the bilabials it is the implosive which is treated as the marked member of the pair (and hence symbolised with the digraph <bh>), whereas for the dentals/alveolars it is the voiced dental stop which is treated as marked, and hence represented with <dh>. Long vowels are written as double vowels. In spontaneously written language, the palatalised allophones of /t/ and /d/ are often written as <c> and <j> respectively, under the influence of Indonesian (e.g. *focu* 'head') The fact that they are only allophones is readily acknowledged by native speakers when this is pointed out, and their unitary writing as <t> hardly poses problems. In the standardised orthography for the southern dialect, the glottal stop is not written in initial position, only intervocalically. However, for linguistic reasons it will be written in this article in initial position too, as in the last line of the example below.

| **phonetic** | **phonemic** | **orthographic** | **gloss** |
|---|---|---|---|
| ['fitʃu] | /fitu/ | *fitu* | 'seven' |
| ['tʲʰu] | /t̪uu/ | *tuu* | 'knee' |
| ['ɓadʒu] | /ɓad̪u/ | *bhadhu* | 'shirt' |
| ['dʒini] | /d̪ini/ | *dhini* | 'evil spirit' |
| [ɓihi'dʒiŋi] | /ɓihid̪iŋi/ | *bhihidhingi* | 'mock' |
| ['ɗaʔu] | /ɗaʔu/ | *da'u* | 'dog' |
| ['ʔatɔ] | /ʔato/ | *'ato* | 'roof' |

## 2.4 Phonological differences between north and south Muna

The following section is partly based on van den Berg (1991a), though I adopt a purely synchronic perspective and mention only in passing the historical implications for Proto Muna.

The three main phonological differences between the two dialects are as follows:

a. south /h/ for north /r/;
b. south glottal stop for north /gh/, a voiced uvular fricative;
c. south glottal stop for north /h/.

These are all treated in turn below, followed by discussion of a few more minor differences.

### *2.4.1 South /h/ for north /r/*

The sound correspondence between /h/ and /r/ is very conspicuous and is one of the main indicators of the southern dialect complex. Notice the following pairs:

| **south** | **north** | **gloss** |
|---|---|---|
| *bhahi* | *bhari* | 'many' |
| *buhi* | *buri* | 'write' |
| *dahu* | *daru* | 'starfruit' |
| *ehe* | *ere* | 'stand' |
| *hambi* | *rambi* | 'hit' |
| *hobhine* | *robhine* | 'woman' |
| *hoo* | *roo* | 'leaf' |
| *kihi* | *kiri* | 'thorn' |
| *sohihi* | *soriri* | 'side' |
| *tuhu* | *turu* | 'drop (n)' |
| *woha* | *wora* | 'see' |

This sound correspondence is completely regular in native vocabulary and goes back to Proto Muna **r*. The same regular correspondence is also found in many loan words from Indonesian/Malay, as in the following list, with the Malay source provided in brackets following the gloss:

| **south** | **north** | **gloss** | |
|---|---|---|---|
| *bhihita* | *bhirita* | 'news' | *(berita)* |
| *gaha* | *[ghohia]* | 'salt' | *(garam)* |
| *husa* | *rusa* | 'deer' | *(rusa)* |
| *kahadhaa* | *karadhaa* | 'work' | *(kerja)* |
| *kahatasi* | *karatasi* | 'paper' | *(kertas)* |
| *kahumbau* | *karambau* | 'buffalo' | *(kerbau)* |
| *nahakaa* | *narakaa* | 'hell' | *(neraka)* |
| *po-guhu* | *po-guru* | 'learn' | *(guru)* |

However, there are many instances where /r/ is retained in loan words. In these instances there is no difference between the two dialects in this respect.

| **south** | **north** | **gloss** | |
|---|---|---|---|
| *'angguru* | *angguru* | 'wine' | *(anggur)* |
| *banara* | *banara* | 'true' | *(benar)* |
| *guru* | *guru* | 'teacher' | *(guru)* |
| *haragaa* | *haragaa* | 'price' | *(harga)* |
| *parasaea* | *parasaea* | 'believe' | *(percaya)* |
| *piara* | *piara* | 'raise, keep' | *(piara)* |
| *ra'eati* | *raeati* | 'populace' | *(rakyat)* |
| *rampasi* | *rampasi* | 'rob' | *(rampas)* |
| *rohi* | *rohi* | 'spirit' | *(roh)* |
| *roti* | *roti* | 'bread' | *(roti)* |
| *surugaa* | *surugaa* | 'heaven' | *(surga)* |

There are also words with /r/ for which the etymology is as yet unclear, although presumably they are loan words too, either from the northern dialect, from Dutch (Du), or from the neighbouring language of Wolio (Wol).

| **south** | **north** | **gloss** | |
|---|---|---|---|
| *boro* | *boro* | 'injection' | (Du *boor*) |
| *burukou* | *wurukou* | 'wood pigeon' | |
| *dharabisi* | *dharabisi* | 'moustache' | |
| *haroa* | *haroa* | 'ritual meal' | (Wol *haroa*) |
| *here* | – | 'very angry' | |
| *ngara* | *ngara* | 'fed up, bored' | |
| *ntara* | *ntara* | 'hold out, endure' | |
| *rapo-rapo* | *rapo-rapo* | 'peanuts' | (Wol *rapo-rapo*) |

It is unclear why these words with /r/ have resisted the change to /h/. A theory of lexical diffusion might suggest that these words remained untouched by the weakening rule which was clearly still operative when loan words entered the lexicon, though the question remains why. Alternatively one can speculate that these words were borrowed at a time when the sound change /r/ –> /h/ had already completed its course. Certainly loans from Indonesian during the last fifty years all retain /r/: *garedha* 'church', *radio* 'radio' etc. This last alternative seems the more likely scenario, although there is the extra factor of influence from Indonesian through education. The word *guru* 'teacher', for instance, has the variant *guhu*, and I suspect the general use of the Indonesian word *guru* may have played a role in creating the variant, which now seems to be pushing out the older *guhu*. Interestingly, the derived word *po-guhu* 'learn' does not have a variant with /r/, presumably because that derivation does not exist in Indonesian and hence retained its 'old' pronunciation with /h/. A similar case is *kahadhaa* 'work', which is sometimes pronounced as *karadhaa* in the south. This explanation runs into problems with *rapo-rapo* 'peanuts', which has a variant *hapo-hapo* used by the very elderly. Since there is no Indonesian target for this etymon, it remains unclear why there is variation.

### *2.4.2 South glottal for north /gh/, a voiced uvular fricative*

The northern voiced uvular fricative /ʁ/ (written as <gh>) corresponds to a glottal stop in the south. This sound correspondence goes back to Proto Muna **q*, which in turn reflects PAN **q*, probably a uvular stop.

| **south** | **north** | **gloss** |
|---|---|---|
| *'ate* | *ghate* | 'liver' |
| *'efi* | *ghefi* | 'lime' |
| *'ito* | *ghito* | 'black' |
| *'oho* | *ghoro* | 'throw away' |
| *'ule* | *ghule* | 'snake' |
| *bu'ou* | *bughou* | 'new' |
| *foho'u* | *foroghu* | 'drink' |
| *ha'a* | *ragha* | 'branch' |
| *mo'ane* | *moghane* | 'man' |
| *tu'a* | *tugha* | 'hard' |

### *2.4.3 South glottal for north /h/*

Some examples of this contrast are:

| **south** | **north** | **gloss** |
|---|---|---|
| *'ali* | *hali* | 'expensive' |
| *'a'u* | *ghahu* | 'attic' |
| *'ewi* | *hewi* | 'sow, scatter' |
| *'ula* | *hula* | 'face' |
| *da'u* | *dahu* | 'dog' |
| *la'ae* | *lahae* | 'who' |
| *pu'e* | *puhe* | 'navel' |
| *sa'o* | *saho* | 'rafter' |
| *ta'a* | *taha* | 'ripe, cooked' |
| *te'i* | *tehi* | 'sea; afraid' |

This correspondence goes back to Proto Muna **h*, but notice that /h/ is retained in loan words and one pronoun:

| **south** | **north** | **gloss** | |
|---|---|---|---|
| *'ahadhi* | *ahadhi* | 'week' | (BI *ahad*) |
| *halia* | *halia* | 'restless, nervous' | |
| *handu* | *handu* | 'towel' | (Du *handdoek*) |
| *hela* | *hela* | 'sail; pull' | (Wol *hela*) |
| *(i)hintu* | *hintu* | 'you (sg)' | |
| *hodha* | *hodha* | 'attempt to persuade' | |
| *kahitela* | *kahitela* | 'maize' | |

The word for 'maize' is derived from Malay *kasitela* 'Spanish/Portuguese', which is usually found as a modifier for various new crops, but has become a free noun in many modern languages (cf. Modern Indonesian *ketela*).

Because of these three very regular sound correspondences, the two dialects sound very differently. Compare the following two sentences containing identical lexical items:

North: *Norato rambi raa mata nokalamo noforoghu oe we ghabu.*

South: *Nohato hambi haa mata nokalamo nofoho'u 'oe we 'abu.*
'He arrived at two o'clock and went to drink water in the kitchen.'

Few northerners take the trouble to learn the southern dialect, but if they do it gives them little phonetic difficulty. Southerners learning the northern dialect, on the other hand, have a dual disadvantage linguistically. Firstly many of them have trouble pronouncing the uvular fricative /gh/, which is so distinctive of the north. They often end up pronouncing it as the velar plosive /g/. Secondly they know they have to replace their glottal stops with either /gh/ or /h/, but often don't know which one to choose. Knowledge that northern /r/ has to be substituted for southern /h/ also leads to hypercorrective forms like *karitela* for correct *kahitela* 'maize'.

It is worth mentioning that the three isoglosses which have been discussed in some detail all represent innovations in the southern dialect. Wider comparative evidence clearly suggests that the northern dialect has retained the original pronunciation (/r/, /gh/ and /h/), whereas the southern dialect has undergone the innovations. This means that the central and prestigious dialect area actually represents a relic area, whereas the innovations have occurred in a more peripheral location with low prestige.

Since change tends to spread from centres of prestige, this is surprising and requires further research.

### *2.4.4 Creaky voice*

In some cases the southern dialects display non-phonemic creaky voice on vowels. This was almost always observed in instances where the word contained both a glottal stop and an implosive stop. Examples:

| | | |
|---|---|---|
| *da'u* | [ɗa̰ʔṵ] | 'dog' |
| *bho'a* | [ɓɔ̰ʔa̰] | 'split, chop' |

### *2.4.5 The phonetics of initial zero*

In the north vowel-initial words are often pronounced with a non-phonemic glottal onset:

| | | | | |
|---|---|---|---|---|
| *ina* | [ina] | ~ | [ʔina] | 'mother' |
| *uli* | [uli] | ~ | [ʔuli] | 'rudder' |

Since the glottal stop is phonemic in the south, there is a contrast between initial glottal and initial zero, as shown in the following pairs:

| **south** | **north** | **gloss** |
|---|---|---|
| *'ule* | *ghule* | 'snake' |
| *ule* | *ule* | 'wave, swing' |
| *'ato* | *ghato* | 'roof' |
| *ato* | *ato* | 'accompany' |

| | | |
|---|---|---|
| *'efī* | *ghefī* | 'lime' |
| *efī* | *ifī* | 'fire' |
| *'owa* | *ghowa* | 'near (south); under (north)' |
| *owa* | *owa* | 'bring' |

What is interesting about initial zero in such cases in the south is that phonetically the onset of the vowel is not zero. There is definitely something to be heard there, but the precise nature of the onset is rather hard to pin down, and the following description must remain impressionistic until detailed acoustic analysis can be carried out. Before the voicing starts the mouth and lips have already taken the shape of the next vowel and there is what may be called a weak voiceless onset or possibly a very weak whisper leading up to the real vowel. The whole process seems to be aimed at a gradual energy build-up to the vowel, avoiding the abrupt start of the glottal stop and also avoiding the friction of initial /h/. The resulting initial vowel is somewhat longer than a vowel preceded by a consonant, especially in stressed position, as exemplified below.

| | | |
|---|---|---|
| *olu* | [ɔ̥ɔlu] | 'cloud' |
| *efī* | [ɛ̥ɛfi] | 'fire' |

Initial zero is not represented in spelling, although some native speakers spontaneously write a long vowel in such cases, e.g. *oolu* and *eefī*.

### *2.4.6 Irregular glottal correspondence with zero*

In addition to the regular sound correspondences /gh/ – /ʔ/ and /h/ – /ʔ/, there are occasional glottals in the south which correspond to zero in the north. This phenomenon is most apparent in loan words from Malay/Indonesian, where the following cases can be distinguished: (1) loanwords with initial vowels are preceded by a glottal; (2) a glottal stop in Malay/Indonesian (itself an allophone of /k/ in postvocalic position or a non-phonemic sound splitting up sequences of like vowels) or from Wolio is retained in the south, but lost in the north; and (3) residual cases. These are illustrated below (loans are from Malay/Indonesian, unless otherwise indicated):

| | **south** | **north** | **gloss** | |
|---|---|---|---|---|
| (1) | *'aki* | *aki* | 'car battery' | *(aki,* Du *accu)* |
| | *'ahadhi* | *ahadhi* | 'week' | *(ahad)* |
| | *'akala* | *akala* | 'mind' | *(akal)* |
| | *'amaha* | *amaha* | 'anger' | *(amarah)* |
| | *'amponi* | *amponi* | 'forgive' | *(ampun)* |
| | *'imani* | *imani* | 'faith' | *(iman)* |
| | *'inawa* | *inawa* | 'spirit, soul' | *(*Wol *inawa)* |
| | *'oli* | *oli* | 'lubricating oil' | *(oli,* Du *olie)* |
| | *'udhi* | *udhi* | 'test' | *(uji)* |
| | *'umuhu* | *omuru* | 'age' | *(umur)* |
| (2) | *dho'a* | *dhoa* | 'ritual prayer' | *(doa)* |
| | *fe'ili* | *feili* | 'character' | *(*Wol *fe'ili)* |
| | *ma'ana* | *maana* | 'meaning' | *(makna)* |
| | *ra'eati* | *raeati* | 'populace' | *(rakyat)* |
| | *Lahata'ala* | *Lahataala* | 'God' | *(Allah taala)* |

(3) *mala'ekati* *malaekati* 'angel' *(malaikat)*

Other cases of irregular initial glottals include the following:

Most function words, such as demonstratives, conjunctions and a pronominal prefix, are pronounced with an initial glottal stop:

| **south** | **north** | **gloss** |
|---|---|---|
| *'aini* | *aini* | 'this' |
| *'aitu* | *aitu* | 'that (near addressee)' |
| *'ampa* | *ampa* | 'until' |
| *'ana'a* | *anagha* | 'that (invisible) |
| *'ane* | *ane* | 'if' |
| *'a-* | *a-* | 'first person singular subject' |
| *'o* | *o* | article preceding nouns |

The free personal pronouns show variation. In isolation and in clause-initial position there is more likely to be a glottal; but following a verb or a preposition the form without the glottal is more common:

| **south** | **north** | **gloss** |
|---|---|---|
| *'anoa ~ anoa* | *anoa* | 'he, she, it' |
| *'andoa ~ andoa* | *andoa* | 'they' |
| *'idi ~ idi* | *idi* | 'I' |
| *'inodi ~ inodi* | *inodi* | 'I' |
| *'insaodi ~ (i)nsaodi* | *insaidi* | 'we (exclusive)' |
| *'intaodi ~ (i)ntaodi* | *intaidi* | 'we two (inclusive)' |

I know of no good explanation to account for the initial glottal in these cases. If we assume that Proto Muna did not have initial glottals but that a non-phonemic glottal could be added to vowel-initial words (as is still the case in the north), then it seems that in the case of some function words this option has become obligatory. With the emergence of /ʔ/ as a full-fledged phoneme in the south, the onset has become phonemicised in a number of cases. Why this development should have been limited to certain function words remains an open question.

Finally there are some irregular cases not involving loan words. Again, it is unclear why a glottal has developed here in the south (assuming that these words did not have an initial Proto Muna **q*):

| **south** | **north** | **gloss** |
|---|---|---|
| *'aa* | *aa* | 'waist' |
| *'i'i* | *ihi* | 'flesh' |
| *'oe* | *oe* | 'water' |
| *'ene* | *ene* | 'pick up' |
| *'uta* | *uta* | 'pick (fruit)' |

### *2.4.7 Intrusive glottal*

When a sequence of three vowels is created by suffixation, a glottal stop is inserted after the second vowel. Common suffixes causing this glottal insertion are the direct object suffix *-e* 'him, her, it' and the indirect object suffixes *-ane*, *-angko*, *-anda* and other vowel-initial

suffixes such as *-omu* 'plural marker on verbs' and *-ana* 'let us (dual)'. In the north such a sequence of three vowels is perfectly acceptable.

Compare the following examples: in (a) the vowel-initial suffix is simply affixed to the verbal base, but in (b) the southern dialect inserts a glottal stop following a sequence of like or unlike vowels (hyphens mark morpheme breaks):

| | **south** | **north** | **gloss** |
|---|---|---|---|
| (a) | *a-woha-e* | *a-wora-e* | 'I see it' |
| | *no-hako-e* | *no-rako-e* | 'he caught him' |
| | *me-'ondo-omu* | *me-ghondo-omu* | 'look!' |
| (b) | *a-humaa-'e* | *a-fumaa-e* | 'I eat it' |
| | *a-fe'ulai-'e* | *a-fehulai-e* | 'I remember it' |
| | *do-tei-'e* | *do-tei-e* | 'they put him' |
| | *no-sambahea-'angko* | *no-sambahea-angko* | 'he prayed for you' |
| | *do-pogau-'omu* | *do-pogau-umu* | 'we (inclusive) speak' |
| | *do-pesua-'ana* | *do-pesua-ana* | 'let us two enter' |

Glottal insertion applies only to sequences of three vowels created by suffixation. In the following three cases glottal insertion does not take place:

1. Monomorphemic words containing three vowels (of which there are relatively few; only in one case a glottal is optional):

| | |
|---|---|
| *daoa* | 'market' |
| *siua* | 'nine' |
| *buea* | 'crocodile' |
| *waea* | 'small bat' |
| *kaue* | 'long wave' |
| *koie ~ koe ~ ko* | 'don't' |
| *paie ~ pae ~ pe* | 'will not' |
| *bheau ~ bhea'u* | 'candle-nut tree' |

2. Prefixes:

| | |
|---|---|
| *tae-afa* | 'what shall we (excl) do?' |
| *koe-ate* | 'way of life' |
| *nao-ambano* | 'he/she will be ashamed' |

3. A very interesting situation as regards the glottal stop arises when the next element is the clitic *-a*. This clitic, the meaning of which is rather elusive, occurs following negators, in exclamations and questions and phrase-finally as a pausal clitic in traditional speaking styles. When this clitic follows the vowel /a/ glottal insertion is optional. Compare the following examples (only from the south):

| | | |
|---|---|---|
| *pae a-k[um]ala-'a* | *~ akumalaa* | 'I won't go' |
| *miina ao-ha-'a* | *~ aohaa* | 'I didn't see' |
| *miina a-s[um]ikola-'a* | *~ asumikolaa* | 'I don't go to school' |

When the clitic follows a sequence of two (or more) vowels of which the last one is /a/, glottal insertion is obligatory:

| | |
|---|---|
| *miina da-tolea-'a* | 'they did not fit' |
| *pae na-[m]esua-'a* | 'he will not go in' |

| | |
|---|---|
| *miina da-[m]arasaea-'a* | 'they don't believe' |
| *miina da-po-me-taa-'a* | 'they don't have a good relationship' |

But when the clitic follows a sequence of two vowels of which the second one is not /a/, glottal insertion does not occur:

| | |
|---|---|
| *miina bhe ka-peo-a* | 'it did not go out (of fire)' |
| *pae na-ko-diu-a* | 'he will remain silent' |
| *miina na-pu-gau-a* | 'she did not speak' |
| *koe sobae-a* | 'don't try it' |

However, some glottals tend to get dropped, both intrusive glottals (especially with *-ane* following the high vowels /i, u/) and inherent glottals such as in the applicative suffix *-'ao* (*-ghoo* in the north):

| | | |
|---|---|---|
| *do-po-bhai-'anda* | ~ *dopobhaianda* | 'they accompanied him' |
| *no-sumpui-'ane* | ~ *nosumpuiane* | 'he met him with s.t.' |
| *a-kadiu-'angko* | ~ *akadiuangko* | 'I bathed (him) for you' |
| *mai-'ao* | ~ *maiao* | 'from' |
| *kampoho'u'a* | ~ *kampohou'a* | 'drinking vessel' |

In conclusion, the origins of the southern glottal stop are manifold. In the first place it reflects Proto Muna **q*, in the second place Proto Muna **h*, in the third place it occurs in loans to replace a medial glottal or an initial non-phonemic glottal, in the fourth place many function words have developed an initial glottal, and in the fifth place (the 'intrusive glottal') it represents a word constraint to break up certain vowel sequences. In a number of cases the history of the glottal remains unclear.

### *2.4.8 Sporadic changes*

1. In some words a medial /l/ is dropped in the south; this is observed in the two frequently used verbs *late* 'live, stay' and *lodo* 'stay' (notice that these verbs are always inflected and never occur in their root form; the southern root forms are therefore *ate* and *odo*). L-deletion is also found in a preposition and in one noun preceded by a preposition:

| **south** | **north** | **gloss** |
|---|---|---|
| *ne-ate* | *ne-late* | 'he lives' |
| *kae-ate-'a* | *kae-late-ha* | 'place of living' |
| *no-odo* | *no-lodo* | 'he sleeps' |
| *ka-odo-'a* | *kao-lodo-ha* | 'sleeping place, bed' |
| *weo* | *welo* | 'in, inside' (shortening of *we lalo*) |
| *se ambu* | *we lambu* | 'at home' |
| *buusi ~ bulusi* | *[ghubhe]* | 'floor beams' |

L-deletion does not take place in initial position and also many other potential words fail to undergo it:

| **south** | **north** | **gloss** |
|---|---|---|
| *lambuku* | **ambuku* | 'my house' |
| *nolabhi* | **noabhi* | 'it is more' |
| *nelangke* | **neangke* | 'it is tall' |
| *te lani* | **te ani* | 'in the sky' |

2. Following /h/, which corresponds to northern /r/, the southern dialect has lowered the vowel /u/ to /o/ in a few cases:

| **south** | **north** | **gloss** |
|---|---|---|
| *hobu* | *rubu* | 'small' |
| *hodua* | *rudua* | 'two (people)' |
| *kahoku* | *karuku* | 'grass, weeds' |

But not in the following cases:

| | | |
|---|---|---|
| *hunsa* | *runsa* | 'leave, divorce' |
| *kahumba* | *karumba* | 'awl' |

3. Spontaneous prenasalisation:

| **south** | **north** | |
|---|---|---|
| *kamongkula ~ kamungkula* | *kamokula* | 'old, parent' |
| *kamontu'a ~ kamuntu'a* | *kamotugha* | 'woods, forest' |
| *'ondula* | *ghonula* | 'shadow' |

4. Others, involving various consonant and vowel alternations:

| **south** | **north** | |
|---|---|---|
| *'aini ~ 'ane* | *ane* | 'if' |
| *'ampa ~ 'ompa* | *ampa* | 'until' |
| *awua* | *awa* | 'grandchild, grandparent' |
| *bhae* | *pae* | 'husked rice' |
| (but both dialects *pae* 'rice in the field') | | |
| *bhaha'i* | *bhahi* | 'whether, or' |
| *bhahitie* | *bhasitie* | 'brother, family' |
| *bundolo* | *bunsolo* | 'eye (rude term)' |
| *daanu(mo)* | *daano* | 'indeed' |
| *dhoi* | *doi* | 'money' |
| *foina'u* | *fenaghu* | 'to teach, advise' |
| *gahuha* | *galura* | 'heavy rain with strong wind' |
| *humaa* | *fumaa* | 'to eat' |
| *kabeanga* | *kabilanga* | 'just like' |
| *kabhao-bhaono* | *bhaa-bhaano* | 'first' |
| *kakuhua* | *kakurao* | 'to crow' |
| *kane'o* | *niho* | 'just now' |
| *kanduula* | *kandulua* | 'pillow' |
| *kaowu* | *kaawu* | 'only' |
| *kodo'o ~ kudo'o* | *kodoho* | 'far' |
| *ladhima* | *adhima* | 'charm, amulet' |
| *lense* | *lensi* | 'to loose, untie' |
| *mongiwa* | *moniwa* | 'shark' |
| *no'ua ~ nu'ua* | *nuhua* | 'pitcher' |
| *poindalo* | *pindalo* | 'to wish' |
| *sadhia ~ saodhia* | *sadhia* | 'always' |
| *se-hewu* | *se-riwu* | 'one thousand' |
| *sia'e ~ sea'e* | *siaghe* | 'too' |
| *sigao'ano* | *sigaahano* | 'other' |

| | | |
|---|---|---|
| *sikaea* | *sirikaea* | 'sweet-sop fruit' |
| *susu ~ tusu* | *tusu* | 'to point' |
| *tantusu ~ ntasusu* | *tantusu* | 'index finger' |
| *tinangke* | *kinante* | 'while, as long as' (metathesis) |

5. The verb 'to give' seems to have undergone obligatory vowel shortening, with the base either *wa* or *wa'a*. The northern form is *waa,* with optional shortening to *wa* when an indirect object suffix follows. Some of the more common forms of this verb are as follows:

| **south** | **north** | |
|---|---|---|
| *no-wa'a-kanau* | *no-wa(a)-kanau* | 'he gave me' |
| *no-wa-'angko* | *a-wa(a)-angko* | 'I gave you' |
| *no-wa-'ane* | *no-wa(a)-ane* | 'he gave her' |
| *no-wa-'anda* | *no-wa(a)-anda* | 'he gave them' |
| *no-fo-wa-'ao* | *no-fo-waa-ghoo* | 'he gave us (incl)' |

## 3 Morphology

Although the most conspicuous areas of divergence between the south and the north are in phonology and lexicon, this does not mean that the morphology and syntax of both dialects are identical. Indeed, the morphology of the south differs from the north in some important aspects, to which we will turn now. Within the scope of this article it is impossible to present a full-fledged overview of southern Muna morphology. The reader will be assumed to be familiar with northern Muna morphology (as described in van den Berg 1989) and hence only major points of divergence will be noted here.

### 3.1 Free pronouns

In the following chart hyphens mark morpheme breaks; optional initial glottals in the south have not been indicated; the Proto Muna reconstructions are from van den Berg (1991a).

| | | **north** | **south** | **Proto Muna** |
|---|---|---|---|---|
| sg | 1 | *inodi, idi* | *inodi, o idi* | **inodi* |
| | 2 | *(i)hintu* | *hintu* | **(i)sintu* |
| | 2pol | *intaidi* | – | |
| | 3 | *anoa* | *anoa* | **anoa* |
| du | 1inc | *intaidi* | *(i)ntaidi/(i)ntaodi* | **intaodi* |
| pl | 1inc | *intaidi-imu* | *(i)ntaidi-omu/(i)ntaodi-omu* | |
| | 1ex | *insaidi* | *(i)nsaodi* | **incamoodi* |
| | 2 | *(i)hintu-umu* | *hinti-miu* | |
| | 2pol | *intaidiimu* | – | |
| | 3 | *andoa* | *andoa* | **andoa* |

Comments on the free pronouns:

- Polite forms are not in use in the south. This is probably a reflection of the fact that there is no nobility class in the south which uses the titles *La Ode* (for men) and *Wa Ode* (for women).
- *idi* is a reduced form of *inodi*, with the article *o* obligatorily present. In writing, this form is often written as one word: *oidi*. Very freqently the form *idia* is heard, containing the pausal clitic *-a*.
- *(i)ntaodi* 'we dual incl.' is the older form, while *(i)ntaidi* (which has undergone vowel assimilation) is very likely dialectal influence from the north. This variation is much less for *(i)nsaodi*, although southern Muna speakers are heard to use the northern form *insaidi.*
- The forms with initial *i-* are more common clause-initially and as free-standing forms. The short forms *ntaodi* and *nsaodi* occur mainly after prepositions, e.g. *ne nsaodi* 'to us', although examples of the short forms clause-initially have also been found. Only infrequently is the form *ihintu* 'you (sg)' heard in the south.

## 3.2 Pronominal affixes

- The subject prefixes are almost identical in both dialects, except that for the second person singular the southern dialect has lost the initial *o-*. The three class prefixes corresponding to the northern *o-*, *ome-* and *omo-* are therefore zero, *me-* and *mo-* in the south. Occasionally full forms with *o-* may be heard (with initial glottal, so *'o-*, *'ome-* and *'omo*), but the short forms are definitely the unmarked situation. Examples with two verbs from each class are:

| **south** | **north** | |
|---|---|---|
| *k[um]ala ne 'amai?* | *o-k[um]ala ne hamai?* | 'where are you going?' |
| *pande-mo!* | *o-pande-mo!* | 'you are clever!' |
| *me-ate ne 'ini?* | *ome-late ne ini?* | 'do you live here?' |
| *me-'oli 'ae?* | *ome-gholi hae?* | 'what did you buy?' |
| *mo-odo-mo?* | *omo-lodo-mo?* | 'do you want to sleep?' |
| *mo-ndawu?* | *omo-ndawu?* | 'did you fall?' |

- The northern plural suffix *-Vmu* (where V stands for a vowel copy of the last vowel of the base), which is used for the second person and first person inclusive, corresponds to *-omu* in the south. The northern form *-Vmu* is actually relatively restricted even in the north, and mainly limited to the old subdistrict of Katobu. The remaining northern subdialects also use *-omu*. Following the suffixes *-ko* 'you (object)' and *-nto* 'our (inclusive)' the dialect difference is neutralised.

| **south** | **north** | |
|---|---|---|
| *kala-omu* | *o-kala-amu* | 'you (pl) went' |
| *suli-omu* | *o-suli-imu* | 'you (pl) returned' |
| *woha-kanau-omu?* | *o-wora-kanau-umu?* | 'did you (pl) see me?' |
| *'a-woha-ko-omu* | *a-wora-ko-omu* | 'I saw you (pl)' |
| *ama-nto-omu* | *ama-nto-omu* | 'our (pl incl) father' |

- The order of this plural suffix *-omu* and the perfective suffix *-mo* in the south is the reverse from the north. In the south the perfective precedes, in the north it follows the plural suffix, with subsequent loss of medial /m/ and vowel assimilation.

| | | |
|---|---|---|
| *kala-mo-omu?* | *o-kala-amoo?* (< *kala-amu-mo) | 'have you (pl) already gone? |
| *[m]oni-mo-omu?* | *[m]oni-imoo?* (< *[m]oni-imu-mo) | 'will you (pl) already go up?' |

Intermediate forms without vowel contraction can be observed in the northern subdialect of Kabawo/Tongkuno, which has forms such as *o-kala-omuo* and *o-moni-omuo*.

- The second person plural possessive suffix in the south is *-miu*, corresponding to the northern *-Vmu* (which is used on pronouns, nouns and verbs). This form *-miu* is also found on passive participles (marked by the prefix *ne-*).

| **south** | **north** | |
|---|---|---|
| *lambu-miu* | *lambu-umu* | 'your (pl) house' |
| *kamungkula-miu* | *kamokula-amu* | 'your (pl) parents' |
| *ne-fetingke-miu* | *ne-fetingke-emu* | 'what you (pl) heard' |
| *ne-woha-miu* | *ne-wora-amu* | 'what you (pl) saw' |

Comparative research suggests that the southern forms are more conservative, and that there has been a gradual movement to replace *-miu* by *-omu*. In the north the form *-miu* is used on a limited number of kinship terms for nobility, such as *ai-miu* 'your wife', used when speaking to a nobleman. Again the evidence seems to be that the northern dialect area is the area of innovation.

- The first person inclusive (dual) object suffix is *-kaita* or *-kainta* in the south. The corresponding form in the north is *-kaeta*, but this is only used as a second person polite suffix, e.g. *a-wora-kaeta* 'I saw you (pol)'. As mentioned before, the southern dialect does not use polite pronominal forms, and hence *-kai(n)ta* has retained its original meaning 'us (inclusive)'. In the north the 'gap' is filled by the detransitivising prefix *fo-*, an option which is also available in the south (see van den Berg 1989:71–72 for details).

| **south** | **north** | |
|---|---|---|
| *no-woha-kainta* | – | 'he sees us (du.incl) |
| *no-fo-woha* | *no-fo-wora* | 'he see us/people (incl)' |
| *no-faraluu-kainta-omu* | – | 'he invited us (all)' |

- The dual form *-kai(n)ta* and its plural equivalent *-kai(n)taomu* are also used in polite imperative or adhortative clauses in which the speaker participates in the action (or politely intends to participate), a usage which is parallelled in the north:

| | | |
|---|---|---|
| *suli-kaita* | *suli-kaeta* | 'let us (2) go home' |
| *me-soso-kainta-omu* | *me-soso-kaeta-amu* | 'let us (all) smoke' |

- The first person exclusive object suffix is *-kansami* or *-kainsami*, corresponding to northern Muna *-kasami*.

*no-tofa-kainsami* *no-tofa-kasami* 'he hit us (excl)'

- The sequence of *-kanau* '(for) me' and *-e* 'him/her/it' is sometimes realised as *-kana-'e*:

  *no-'oli-kana-'e* *no-gholi-kanau-e* 'she bought it for me'

- In rapid speech the subject markers may get dropped altogether, especially for third person singular. This usage is frowned upon by more careful speakers. Some examples from written texts, with the missing prefix provided in brackets:

  *'Oleo (no-)ma'o-mo nao-hondo.*
  sun 3sR-near-PERF 3sI-dark
  'It was almost dark.'

  *(No-)mintae do-sawi-mo ne kapala.*
  3sR-light 3pR-go.by-PERF loc boat
  'The next morning they took a boat.'

  *Pokono (no-)kiido na-[m]o-wanu-e fotu-no.*
  in.short 3sR-refuse 3sI-CAUS-get.up-it head-his
  'In short, he refused to raise his head.'

### 3.3 Allomorphy of the infix *-um-*

The allomorphy of the infix *-um-* hardly differs between the two dialects. This infix is used to form the irrealis for verbs of class *a-* and for verbs of class *ae-* with definite objects. The basic form is *-um-*, which changes to *m-* when the verb root begins with a vowel. When the verb root begins with a glottal, the regular infix *-um-* is triggered. The voiceless bilabial phonemes /p/ and /f/ change to /m/, and with initial /b, m, bh/ there is no change. With initial /w/ there is variation between /w/ and /m/, and some high frequency words such as 'eat', 'drink' and 'see' show reduction. The following examples illustrate these processes (S = south; N = north; translations follow the irrealis form).

| | **realis** | **irrealis** | |
|---|---|---|---|
| S | *no-ehe* | *na-[m]ehe* | 'he will stand up' |
| N | *no-ere* | *na-mere* | |
| S | *ne-'oli* | *na-'[um]oli-e* | 'she will buy it' |
| N | *ne-gholi* | *na-gh[um]oli-e* | |
| S | *ne-'amponi* | *na-'[um]amponi-e* | 'he will forgive him' |
| N | *ne-amponi* | *na-[m]amponi-e* | |
| S | *a-wowo'o-e* | *a-wowo'o-e/a-[m]owo'o-e* | 'I will take it' |
| N | *a-wowoho-e* | *a-wowoho-e* | |
| S | *a-foho'u* | *ao-ho'u* | 'I will drink' |
| N | *a-foroghu* | *ao-roghu* | |
| S | *a-humaa* | *ao-maa* | 'I will eat' |
| N | *a-fumaa* | *ao-maa* | |
| S | *a-woha-e* | *a-[m]oha-e/ao-hae* | 'I will see it' |
| N | *a-wora-e* | *a-[m]ora-e/ao-rae* | |

The existential verb differs slightly in the two dialects, both in realis, irrealis and participle form, with the southern irrealis and participle forms displaying an unexpected prenasalised consonant /nd/:

| **south** | **north** | |
|---|---|---|
| *nandoo* | *naandoo* | 'there is' |
| *na-nd[um]andoo* | *na-n[um]aandoo* | 'there will be' |
| *nd[um]andoo-'ano* | *n[um]aandoo-no* | 'which there are' |

### 3.4 Verb classes

In Muna, each verb belongs to one of three classes, which I have called the *a*-class, *ae*-class and *ao*-class, following the first person singular. Across the dialect division, it appears that some verbs have shifted classes or at least show variation, although at this point it cannot be ascertained in which direction the change has taken place. Some examples are:

| | | | |
|---|---|---|---|
| *ambano* | 'ashamed' | N: *a-*; | S: *a-/ao-*. |
| *bhahi* | 'be many' | N: *a-*; | S: *a-/ao-*. |
| *bhala* | 'big' | N: *a-*; | S: *a-/ao-*. |
| *himba* | 'fast' | N: *ae-*; | S: *ae-/ao-*. |
| *kiido* | 'refuse' | N: *a-*; | S: *a-/ao-*. |

### 3.5 The indirect object suffix *-'ao*

The southern equivalent to the northern indirect object suffix *-ghoo* is *-'ao*, which itself is a reflex of an older form **-ako*, ultimately from **-aken*. The primary use of this affix, which may also be called an applicative suffix, is to indicate the presence of an extra NP argument in the clause, which is not the direct object. Its second use is to indicate purpose. It will be glossed as 'IO' for indirect object or 'PURP'. Some examples, with the first line in the southern dialect, the second line in the northern dialect are:

S *'A-mai-'ao we sikola.*
N *A-mai-ghoo we sikola.*
1sR-come-IO loc school
'I come from school.'

S *Na-fo-wa-'ao ka-'osa.*
N *Na-fo-waa-ghoo ka-ghosa.*
3sI-DETR-give-IO NOM-strong
'He will give us strength'

S *Noafa mo-'ae-'ao-'omu?*
N *Noafa omo-ghae-ghoo-omu?*
why 2s-cry-PURP-PLUR
'Why are you (pl) crying?'

S *'o dhoi so 'ae-'oli-'ao kenta*
N *o doi so ae-gholi-ghoo kenta*
ART money for 1sI-buy-PURP fish
'money to buy fish'

In rapid speech the actual shape of this suffix is sometimes realised as *-'oo* with vowel assimilation or even *-'o* with vowel reduction; very occasionally the glottal is dropped altogether.

| | | |
|---|---|---|
| S | *No-pono-'o* | *kuasa.* |
| N | *No-pono-ghoo* | *kuasa.* |
| | 3sR-full-IO | power |
| | 'it is full of power' | |

### 3.6 Different *-Ci* suffixes

So far most of the dialect differences have been fairly regular. In the case of the locative and iterative suffix *-Ci* the differences appear to be idiosyncratic. In each dialect there are many surface forms (*-fi, -ni, -ki, -pi, -si* etc.) which are lexically conditioned. Although most of the suffixes are identical in the two dialects, there is disagreement in a considerable number of cases. Apparently this is a rather unstable area in the language as a whole, adding weight to the observation in van den Berg (1991b:19) that such variation shows the massive reorderings that have taken place during the history of these verbs. At some point these 'thematic' consonants must have occupied root-final position, but with the loss of final consonants in the parent language the original final consonant was reinterpreted as part of the suffix. Examples of the differences:

| **south** | **north** | |
|---|---|---|
| *dhaga-i ~ dhaga-ni* | *dhaga-ni* | 'guard' |
| *late-fi* | *late-ghi* | 'reside in' |
| *longko-wi* | *longko-fi* | 'bend over to get' |
| *lumpa-ki* | *lumpa-ghi* | 'collide with' |
| *ma'ana-ki* | *maana-ghi* | 'interpret, translate' |
| *ngkoha-fi* | *ngkora-ghi* | 'sit on' |
| *'ondo-fi* | *ghondo-hi* | 'look for' |
| *pesua-ki* | *pesua-ghi* | 'enter' |
| *po-kudo'o-ti* | *po-kodoho-pi* | 'far from each other' |
| *sia-pi ~ sia-ti* | *sia-ti* | 'bite (many times)' |
| *tumbu-si* | *tumbu-li* | 'punch (many times)' |

The majority of verbs, however, have identical or corresponding suffixes, such as the following:

| | | |
|---|---|---|
| *hambi-si* | *rambi-si* | 'hit (many times)' |
| *mama-'i* | *mama-ghi* | 'chew a tobacco quid' |
| *suli-ki* | *suli-ki* | 'return to get' |

In some words the south has a *-Ci* suffix which is lacking in the north, as in the word *'owa-ti* 'approach' from *'owa* 'near'. The corresponding locative noun *ghowa* means 'under' in the north, but there is no verb derived from it. Also: *uumbe-ti* 'agree to; say yes to' from *uumbe* 'yes'. Again, the verb is lacking in the north.

### 3.7 Extra *-Cao* suffixes

In northern Muna there is a *-Cao* suffix which mainly indicates violent action, as in *tumbu* 'pound' vs *tumbu-lao* 'plant firmly, embed'. It is used on a limited number of verbs. It appears that in southern Muna this formation is more productive (or has retained greater productivity) as witnessed by the following forms which do not exist in the northern dialect. Notice that the meanings of these formations are quite idiosyncratic, but all indicate some kind of intensified action.

| | | | |
|---|---|---|---|
| *bhake* | 'spread out' | *bhake-sao* | 'spread out many items' |
| *buso* | 'blow' | *buso-hao* | 'blow forcefully' |
| *gego* | 'shake' | *gego-hao* | 'shake violently' |
| *hompu* | 'gather' | *hompu-nao* | 'gather in large quantities' |
| *tale* | 'lay out neatly' | *tale-sao* | 'lay out neatly in large numbers' |
| *tepi* | 'winnow' | *tepi-sao* | 'winnow in large quantities' |

In at least two cases, the corresponding form exists in the north with a different thematic consonant:

| | | | |
|---|---|---|---|
| *kitu* | 'scrub' | S *kitu-fao* | 'scrub violently; attack' |
| | | N *kitu-rao* | |
| *ehe* | 'stand up' | S *ehe-nao* | 'suddenly stand up' |
| | | N *ere-hao* | |

In addition to these verbs, the south also possesses an extra morphological possibility with these intensive verbs. Instead of using *-Cao*, the action can be even further intensified by using the suffix *-Ca'iao*, where the consonant is again lexically determined. The resulting meaning is something like 'do X roughly and angrily'. Some examples are:

| | |
|---|---|
| *hambi* | 'hit' |
| *hambi-tao* | 'fling down' |
| *hambi-ta'iao* | 'fling down roughly and angrily' |
| *kabhambi* | 'throw away angrily' |
| *kabhambi-lao* | 'throw away violently' |
| *kabhambi-la'iao* | 'throw away violently and in disgust' |
| *tumbu* | 'pound, strike' |
| *tumbu-lao* | 'plant firmly in/on the ground, embed' |
| *tumbu-la'iao* | 'plant very firmly in/on the ground' |

### 3.8 Extra circumfix *fisi-/-'a*

Southern Muna has a productive circumfix *fisi-/-'a* which is lacking in the north. It denotes a sudden action, and is often found in combination with the prefix *ta-* (which also indicates a sudden or unexpected action) and the perfective suffix *-mo*. For example:

| | |
|---|---|
| *ta-do-fisi-hato-'a-mo* | 'suddenly they arrived' |
| *ta-no-fisi-limba-'a-mo* | 'suddenly he went outside' |
| *ta-no-fisi-bhote-'a-mo* | 'suddenly it exploded'. |

### 3.9 Kinship circumfix *foko-/-'ao*

The northern kinship circumfix *foko-/-u* corresponds to *foko-/-'ao* in the south. Compare the following pairs:

| **south** | **north** | |
|---|---|---|
| *fokoama'ao* | *fokoamau* | 'uncle' (cf. *ama* 'father') |
| *fokoina'ao* | *fokoinau* | 'aunt' (cf. *ina* 'mother' |
| *fokoana'ao* | *fokoanau* | 'nephew, niece' (cf. *ana* 'child') |
| *fokoawua'ao* | *fokoawau* | 1. 'great-aunt, great-uncle' |
| | | 2. 'great-nephew, great-niece' (cf. *awua* 'grandchild, grandparent') |

### 3.10 Adhortatives

Northern Muna has two adhortatives: *-ana* (dual) and *-mana* (plural). In the south the dual form is identical, but the plural form is *-ana-omu*, which contains the pluralising suffix *-omu* (see §3.2 above). Possibly the northern form *-mana* can be accounted for by assuming a reverse sequence of the affixes with subsequent fusion, i.e. **-omu-ana* > *-muana* > *-mana*. For example:

| **south** | **north** | |
|---|---|---|
| *do-kala-ana* | *do-kala-ana* | 'let us (2) go' |
| *do-suli-ana-omu* | *do-suli-mana* | 'let us (all) go home' |

### 3.11 Stative verb reduplication

It appears that southern Muna has a unique reduplication pattern on stative verbs (roughly corresponding to adjectives in English) which indicates a very high degree.

The reduplicated vowel is longer than a single vowel, but does not have quite the length of a double vowel, possibly because it occurs in an unstressed position in the word. However, it will be written with two vowels to indicate its lengthened nature. The translation refers to the reduplicated form.

| | | |
|---|---|---|
| *no-bhala* | *no-bhoo-bhala* | 'it is very big' |
| *no-bu'ou* | *no-boo-bu'ou* | 'it is very new' |
| *no-dai* | *no-doo-dai* | 'it is very bad' |
| *no-dea* | *no-doo-dea* | 'it is very red' |
| *no-idho* | *no-oo-idho* | 'it is very green' |
| *no-'ito* | *no-'oo-'ito* | 'it is very black' |
| *no-kuni* | *no-koo-kuni* | 'it is very yellow' |
| *ne-langke* | *ne-lee-langke* | 'it is very tall' |
| *no-mangka* | *no-moo-mangka* | 'he is very exhausted' |
| *no-pute* | *no-poo-pute* | 'it is very white' |
| *ne-taa* | *ne-tee-taa* | 'it is very good' |
| *ne-wanta* | *ne-wee-wanta* | 'it is very long' |
| *no-wule* | *no-woo-wule* | 'he is very tired' |

This pattern of reduplication is very interesting. The reduplicated consonant is simply a copy of the first prevocalic consonant of the root, but the long vowel in the reduplicated syllable is not a copy of the root vowel, but a lengthened verb-class vowel. Notice that the first seven examples all have long *oo* in the reduplicated syllable as they are all *ao*-verbs; *langke* 'tall' and *taa* 'good' are *ae*-verbs and hence the long vowel in the reduplicated syllable is *ee*. It seems that all such reduplicated stative verbs belong to either the *ao*-class or the *ae*-class; no examples of the a-class have been recorded. In fact, regular stative verbs which belong to the *a*-class (such as *bu'ou* 'new'), change to the *ao*-class when they are reduplicated, as in the following elicited example:

*a-bu'ou* *ao-boo-bu'ou* 'I am very new'

This means that in the reduplicated syllable only the consonant is reduplicated. The long vowel must be analysed as pre-associated and lexically specified.

### 3.12 Variation of full reduplication

Full reduplication (that is, reduplication of two syllables) is sometimes reduced to a reduplicated consonant followed by one long vowel, which is a copy of the first reduplicated vowel. This pronunciation is merely a variation of the regular pattern and is frowned upon by more careful speakers of the language. Examples are:

| | | | |
|---|---|---|---|
| *kapo-kapoluka* | ~ | *kaa-kapoluka* | 'tortoise' |
| *kapu-kapuna* | ~ | *kaa-kapuna* | 'story' |
| *kambu-kambuhumaino* | ~ | *kaa-kambuhumaino* | 'the very last' |
| *panda-pandano* | ~ | *paa-pandano* | 'at last' |

### 3.13 Participles with *-'a*

Active participles in both dialects are formed with the circumfixes *-um-/-no*, *me-/-no* and *mo-/-no* for the three verb classes respectively. However, it appears that in the south an extra morpheme *-'a* is often inserted, which in the north is used only to indicate a place, a time or a part of a whole. The reasons for this are not yet clear. Examples (from the south only):

| | |
|---|---|
| *wula k[um]undo-'a-no* | 'last month' |
| *kenta mo-hobu-'i-'a-no* | 'small fishes' |
| *sau pata manso-bhongka-'a-no* | 'wood which does not break easily' |
| *amaitu-'ae maraluu-'a-no sepali'a* | 'that is what is most needed' |

## 4 Syntax

In the area of syntax the dialectal differences are almost negligible, but a few are worth mentioning. Almost all the phenomena here have to do with function words.

### 4.1 The article *o*

The elusive article *o* in Muna has a slightly different distribution in the south, where its shape is *'o*. Its main function appears to be identical in both dialects, namely what I have called 'relative syntactic freedom' of nouns. This means *o/'o* occurs preponderantly in single-word clauses, in lists, with nominal predicates, with preverbal subjects and appositions (see van den Berg 1989:102–108 for examples). But it is my impression that it occurs less often clause-initially in the south, although this remains to be quantified. What is clear is that in the south the article can accompany nouns with a possessive suffix, something which is ungrammatical in the north. The following examples are all taken from texts:

| **south** | **north** | |
|---|---|---|
| *'o neano* | **o neano* | 'its name' |
| *'o gurumu* | **o gurumu* | 'your teacher' |
| *'o inano* | **o inano* | 'his mother' |

Another feature of the south is that a few vowel-initial nouns tend to use the article even following prepositions, an environment in which it is least likely to occur. Some of these words and example sentences are:

*Da-s[um]obu 'o efi amaitu.*
3pI-extinguish ART fire that
'We will put out that fire.'

*Ne-ate bhe 'o ai-no.*
3sR-live with ART younger.sibling-his
'He lives with his younger sibling.'

*Ne-'oli gola-no 'o ani.*
3sR-buy sugar-POS ART bee
'She bought honey.'

Notice also the use of the article in the pronoun *o idi* 'I'. The north never uses the article in this case.

### 4.2 Demonstratives and locatives

The south uses a unique demonstrative (with variants) which is completely lacking in the north.

*sa'a ~ sa'itu ~ sa'ituini* 1. 'that (near speaker or just mentioned)'
2. 'just now'

The northern equivalent is *aniini*. Some examples are:

*Kala ne 'amai sa'itu-a?*
go loc where just.now-CL
'Where have you been just now?

*Sa'a? O, 'a-kala we daoa.*
just.now o 1sR-go loc market
'Just now? Oh, I went to the market.'

*Do-tapi-e te wawo-no galampa sa'ituini.*
3pR-stack-it loc top-POS drying.rack just.now
'Then they stack it on top of that drying rack (which I just mentioned).'

Another very striking feature of the south is the use of the suffix *-'ae* in locative and temporal expressions. Whereas the normal northern expression for 'here' is *ne ini* or *na ini*, the south frequently uses *ne 'ini'ae.* In the north the use of *-hae* always has predicative force: *ne inihae* means 'here it is', but not so in the south. *Aitu'aemo* often seems to mean 'that is', followed by an explanation or an example (possibly influenced by Indonesian *yaitu*). Notice the following southern examples, all taken from texts:

*Punda ne 'ini-'ae.*
jump loc this-LOC
'jump here'

*Amba-do mie [m]ande-no ma'ana-no 'aitu-'ae-mo.*
word-POS person clever-A.PART meaning-its that-LOC-PERF

*ka-tohopo-no lalo*
NOM-calm-POS heart
'Wise people said that it meant peace.'

*Ne-tisa-mani 'aitu-'ae-mo 'o 'ai, 'o dhambu bhe kalei.*
P.PART-plant-our that-LOC-PERF ART coconut ART cashew and banana
'What we planted were coconut trees, cashew trees and banana trees.'

## 4.3 The preposition *se*

Although the locative preposition *se* is known in the north, it is infrequent and occupies a marginal position in the total system of prepositions. This preposition is much more frequently used in the south, where it indicates a relatively close location which is more or less on a level with the point of orientation.

*'A-k[um]ala se kaampo.*
1sI-go loc field
'I'm going to my field.'

*'A-kala se kamungkula-no liwu.*
1sR-go loc elder-POS village
'I went to the village elders.'

## 4.4 Questions and question words

On a number of points this is an area of some divergence between the two dialects.

- Some different question words (those not listed here are identical, with regular phonological correspondences):

| **south** | **north** | |
|---|---|---|
| *(e)ndefie(mo)* | *indefie* | 'when (past)' |
| – | *nefiemo* | 'when (past)' |
| *se'ae* | *sehae* | 'how much' (functions as a predicate) |
| *fee* | *sehae* | 'how many' (functions as a modifier preceding classifiers and measure nouns; *sehae* can also function as a verb in the north) |

S *Se'ae* *'oli-no?*
N *Sehae* *gholi-no?*
how.much price-its
'How much is it?'

S *Se'ae* *ka-wanta-no?*
N *Sehae* *ka-wanta-no?*
how.much NOM-long-its'
'How long is it?'

S *Fee* *ta'u?*
N *Sehae* *taghu?*
how.many year?
'How many years?'

S *Fee* *mie-mo* *ana-mu*
how.many person-PERF child-your

N *Do-sehae-mo* *ana-mu?*
3pR-how.many-PERF child-your
'How many children have you got?'

- The south uses an alternative question word *kaa* (or *ka*), probably from the Indonesian question clitic *-kah*. This word (which is rarely if ever used in the north) is very common in the south for forming alternative questions, although it can also occur in yes/no questions without a following alternative, where it signals surprise or anger. I will gloss it as 'or'.

*'O-asi-ane* *kaa* *miina?*
2sR-like-it or not
'Do you like it or not?'

*Mai* *moisa* *ka* *bhe-mo* *sabhangka-mu?*
come alone or be-PERF friend-your
'Did you come alone or do you have a friend?'

*Pae-mo* *[m]o-pansuhu-e* *ka-sikola-mu-a* *kaa?*
FUT.not-PERF CAUS-continue-it NOM-school-your-CL or
'Will you not continue your education?'

*Nandoo* *bhe* *kalei-mu* *kaa?*
exist with banana-you or
'So you have some bananas?'

- A particular strong feature of the south seems to be the use of rhetorical echo questions. In the following pair of utterances the rhetorical echo question signals the hearer's mild disagreement. Speaker A expresses thanks for being allowed to stay overnight after a public vehicle broke down, after which speaker B reacts with a rhetorical echo question.

A: *Tumpu lalo-ku kaasi, ingka 'a-feka-mahasai-ko-mo.*
trunk heart-my poor.thing as.you.know 1sR-CAUS-difficult-you-PERF

B: *Feka-mahasai 'ae-no-no padaa? Bheanea se-paku-no-a*
CAUS-difficult what-its-its EMPH maybe one-time-its-CL

*'a-kahondoa toha idi.*
1sR-night.overtaken again I

A: 'Thank you very much. I'm sorry I'm just making things difficult for you.'

B: 'What do you mean making things difficult? You never know, at some time I may be overtaken by the night as well.'

Another example of mild correction is the following short dialogue:

A: *'A-[m]ansuhu-angko we lambu-mu.*
1sI-continue-you loc house-your
'I'll drive you home'

B: *Ka-kudo'o dua kunae.*
NOM-far also MODAL
'It's long way (so maybe you shouldn't do it).'

A: *Na-se'ae-mo dua ka-kudo'o-no?*
FUT-how.much-PERF also NOM-far-its
'It's not far at all.' (lit: How much is the distance too?)

The following pair showing a rhetorical question is the traditional way of greeting and response when one approaches someone's house and can't see anyone in. This is a unique greeting pattern for the south:

A: *Me-gau-mo-omu 'ae?*
2s-cook-PERF-PLUR what
'What are you (already) cooking?'

B: *Tae-gau 'ae? Foni te lambu.*
1pe-cook what go.up loc house
'What are we cooking? Come up in the house'

## 5 Lexicon

The following word list shows the differences in basic vocabulary. The original 210-word list was used by SIL in language surveys in Sulawesi and is based on the Swadesh 200 list, with some culture-specific modifications. A number of items on the original list have been deleted, either because the Muna equivalents are absent or difficult to find (e.g. 'lake', 'guest', 'bark cloth'), or because the lexical item occurred twice in the list (e.g. *matano oe*

'spring', a combination of *mata* 'eye' and *oe* 'water'). The resulting number of words has therefore been reduced to 197.

Notice that the Indonesian equivalents have been added for convenience, and that the orthography for both dialects is phonemic. For ease of reading I repeat the earlier information here: <bh> is a bilabial implosive, <d> is an alveolar implosive, slightly stronger in the south, <dh> is a voiced dental plosive; <t> is affricated in the south preceding /i/ and /u/. Asterisks refer to notes at the end of the list.

| | English | Indonesian | Muna – south | Muna – north |
|---|---|---|---|---|
| 1 | wood | *kayu* | *sau* | *sau* |
| 2 | tree | *pohon* | *pu'u* | *pughu* |
| 3 | leaf | *daun* | *hoo* | *roo* |
| 4 | root | *akar* | *pahaka* | *paraka* |
| 5 | seed | *biji* | *'onu* | *ghonu* |
| 6 | thorn | *duri* | *kihi* | *kiri* |
| 7 | coconut | *kelapa* | *'ai* | *ghai* |
| 8 | coconut shell | *tempurung* | *ka'abulu* | *kaghabulu* |
| 9 | pandanus | *pandan* | *ponda* | *ponda* |
| 10 | banana | *pisang* | *kalei* | *kalei* |
| 11 | fruit | *buah* | *bhake* | *bhake* |
| 12 | betel | *sirih* | *gili, kahoo** | *gili, karoo** |
| 13 | rattan | *rotan* | *'ue* | *ghue* |
| 14 | bamboo | *bambu* | *pahawata* | *koo, patu** |
| 15 | flower | *bunga* | *kambea* | *kambea* |
| 16 | tall grass | *alang-alang* | *dana* | *dana* |
| 17 | rice (in the field) | *padi* | *pae* | *pae* |
| 18 | husked rice | *beras* | *bhae* | *pae* |
| 19 | cooked rice | *nasi* | *bhae, nasi* | *ghoti** |
| 20 | I | *saya, aku* | *inodi, 'o idi* | *inodi, idi* |
| 21 | you | *engkau, anda* | *hintu* | *(i)hintu* |
| 22 | he, she | *dia, ia* | *anoa* | *anoa* |
| 23 | we (exclusive) | *kami* | *(i)nsaodi* | *insaidi* |
| 24 | we (inclusive) | *kita* | *(i)ntaodi(omu)** | *intaodi(imu)** |
| 25 | they | *mereka* | *andoa* | *andoa* |
| 26 | sun | *matahari* | *'oleo* | *gholeo* |
| 27 | moon | *bulan* | *wula* | *wula* |
| 28 | star | *bintang* | *kulipopo* | *kolipopo* |
| 29 | water | *air* | *'oe* | *oe* |
| 30 | rain | *hujan* | *'use* | *ghuse* |
| 31 | stone | *batu* | *kontu* | *kontu* |
| 32 | sand | *pasir* | *bhone* | *bhone* |
| 33 | land, soil | *tanah* | *wite* | *wite* |
| 34 | cloud | *awan* | *olu* | *olu* |
| 35 | wind | *angin* | *kawea* | *kawea* |
| 36 | sea | *laut* | *te'i* | *tehi* |
| 37 | forest, woods | *hutan* | *kahoku, kamontu'a, ponue** | *karuku, katugha* |
| 38 | mountain | *gunung* | *gunu* | *kabhawo* |

| | English | Indonesian | Muna – south | Muna – north |
|---|---|---|---|---|
| 39 | fire | *api* | *efi* | *ifi* |
| 40 | burn | *bakar* | *de-tunu** | *de-tunu* |
| 41 | smoke | *asap* | *gawu* | *ghumbo* |
| 42 | ashes | *abu* | *'abu* | *ghabu* |
| 43 | hot | *panas* | *no-sodo* | *no-pana* |
| 44 | lukewarm | *hangat* | *no-panangkuku, no-mansodo-nsodo* | *no-panangkuku* |
| 45 | cold | *dingin* | *no-hindi* | *no-rindi* |
| 46 | night | *malam* | *kohondo'a* | *korondoha* |
| 47 | fish | *ikan* | *kenta* | *kenta* |
| 48 | bird | *burung* | *ka'uhi (lumola)** | *manu-manu* |
| 49 | egg | *telur* | *'unteli* | *ghunteli* |
| 50 | dog | *anjing* | *da'u* | *dahu* |
| 51 | bat | *keluang, kalong* | *waea* | *waea* |
| 52 | louse | *kutu* | *otu* | *otu* |
| 53 | mosquito | *nyamuk* | *konunu* | *buroto* |
| 54 | mouse | *tikus* | *wulawo* | *wulawo* |
| 55 | snake | *ular* | *'ule* | *ghule* |
| 56 | horn | *tanduk* | *tandu* | *tandu* |
| 57 | tail | *ekor* | *lensi* | *punda* |
| 58 | black | *hitam* | *no-'ito* | *no-ghito* |
| 59 | white | *putih* | *no-pute* | *no-pute* |
| 60 | red | *merah* | *no-dea* | *no-dea* |
| 61 | yellow | *kuning* | *no-kuni* | *no-kuni* |
| 62 | green | *hijau* | *no-idho* | *no-idho* |
| 63 | one | *satu* | *dise, se-** | *ise, se-** |
| 64 | two | *dua* | *dua, haa-* | *dua, raa-* |
| 65 | three | *tiga* | *tolu, tolu-* | *tolu, tolu-* |
| 66 | four | *empat* | *paa, fato-* | *paa, fato-* |
| 67 | five | *lima* | *dima, lima-* | *lima/dima, lima-* |
| 68 | six | *enam* | *noo, nomo-* | *noo, nomo-* |
| 69 | seven | *tujuh* | *pitu, fitu-* | *pitu, fitu-* |
| 70 | eight | *delapan* | *alu, alu-* | *oalu, alu-* |
| 71 | nine | *sembilan* | *siua, siua-* | *siua, siua-* |
| 72 | ten | *sepuluh* | *ompulu* | *ompulu* |
| 73 | twenty | *dua puluh* | *haa-fulu* | *raa-fulu* |
| 74 | hundred | *seratus* | *mo'ono* | *moghono* |
| 75 | thousand | *seribu* | *se-hewu* | *se-riwu* |
| 76 | all | *semua* | *ko-sa-sawi-'ae* | *ko-si-bhari-bhari-hae* |
| 77 | many | *banyak* | *no-bhahi* | *no-bhari* |
| 78 | big | *besar* | *no-bhala* | *no-bhala* |
| 79 | small | *kecil* | *no-hobu* | *no-rubu* |
| 80 | long | *panjang* | *ne-wanta* | *ne-wanta* |
| 81 | short | *pendek* | *ne-'obu, no-mbubuku* | *no-ngkubu* |
| 82 | near | *dekat* | *no-ma'o* | *no-maho* |

| | English | Indonesian | Muna – south | Muna – north |
|---|---|---|---|---|
| 83 | far | *jauh* | *no-kudo'o* | *no-kodoho* |
| 84 | full | *penuh* | *no-pono* | *no-pono* |
| 85 | new | *baru* | *no-bu'ou* | *no-bughou* |
| 86 | good | *baik* | *ne-taa* | *ne-taa* |
| 87 | round | *bulat* | *ne-'onu* | *ne-ngkonu* |
| 88 | dry | *kering* | *no-kele* | *no-kele* |
| 89 | to dry | *menjemur* | *de-'oleo* | *de-gholeo* |
| 90 | not | *tidak* | *miina* | *miina* |
| 91 | this | *ini* | *'aini* | *aini* |
| 92 | that | *itu* | *'aitu* | *aitu* |
| 93 | here | *di sini* | *ne 'ini* | *ne ini* |
| 94 | there | *di situ* | *ne 'itu* | *ne itu* |
| 95 | inside | *di dalam* | *we lalo* | *we lalo* |
| 96 | over, on top of | *di atas* | *te wawo* | *te wawo* |
| 97 | outside | *di luar* | *we sembali* | *we sembali* |
| 98 | under | *di bawah* | *we panda* | *we ghowa, we panda* |
| 99 | before | *di depan* | *te wise* | *te wise* |
| 100 | behind | *de belakang* | *we kundo* | *we kundo* |
| 101 | east | *timur* | *timbu, mata'oleo** | *mata gholeo** |
| 102 | west | *barat* | *bhaha, kansoopa** | *kansoopa** |
| 103 | skin | *kulit* | *kuli* | *kuli* |
| 104 | flesh | *daging* | *'i'i* | *ihi* |
| 105 | fat (n) | *lemak* | *tabha* | *tabha* |
| 106 | blood | *darah* | *hea* | *rea* |
| 107 | heart | *jantung* | *bhake* | *bhake* |
| 108 | liver | *hati* | *'ate* | *ghate* |
| 109 | bone | *tulang* | *buku* | *buku* |
| 110 | body hair | *bulu* | *wulu* | *wulu* |
| 111 | head | *kepala* | *fotu* | *fotu* |
| 112 | face | *muka* | *'ula* | *hula* |
| 113 | eye | *mata* | *mata* | *mata* |
| 114 | nose | *hidung* | *nee* | *nee* |
| 115 | mouth | *mulut* | *wobha* | *wobha* |
| 116 | lip | *bibir* | *wiwi* | *wiwi* |
| 117 | tooth | *gigi* | *wangka* | *wangka* |
| 118 | tongue | *lidah* | *lela* | *lela* |
| 119 | ear | *telinga* | *tingala* | *pongke* |
| 120 | neck | *leher* | *wu'u* | *wughu* |
| 121 | breast | *buah dada* | *titi* | *titi* |
| 122 | stomach | *perut* | *handa* | *taghi* |
| 123 | hand, arm | *tangan* | *lima* | *lima* |
| 124 | nail | *kuku* | *konisi* | *konisi* |
| 125 | foot, leg | *kaki* | *'a'e* | *ghaghe* |
| 126 | knee | *lutut* | *tuu* | *tuu* |
| 127 | urine | *air kencing* | *totolea, 'o'oha** | *ghoghora* |
| 128 | faeces | *tahi* | *ta'i* | *kaedeha* |
| 129 | name | *nama* | *nea* | *nea* |

| | English | Indonesian | Muna – south | Muna – north |
|---|---|---|---|---|
| 130 | person | *orang* | *mie* | *mie* |
| 131 | deaf | *tuli* | *no-pongke* | *no-pongke* |
| 132 | blind | *buta* | *no-bunto* | *no-pilo* |
| 133 | man | *laki-laki* | *mo'ane* | *moghane* |
| 134 | woman | *perempuan* | *hobhine* | *robhine* |
| 135 | father | *ayah, bapak* | *ama* | *ama* |
| 136 | mother | *ibu* | *ina* | *ina* |
| 137 | child | *anak* | *ana, ana'i* | *ana, anahi* |
| 138 | youngest child | *anak bungsu* | *ka'epu* | *kahepu, kampufu* |
| 139 | grandchild | *cucu* | *awua* | *awa* |
| 140 | older sibling | *kakak* | *isa* | *isa* |
| 141 | younger sibling | *adik* | *ai* | *ai* |
| 142 | uncle | *paman* | *fokoama'ao* | *fokoamau* |
| 143 | aunt | *bibi* | *fokoina'ao* | *fokoinau* |
| 144 | ancestors | *nenek moyang* | *awua nsubhe* | *awa fitu tapino* |
| 145 | friend | *teman, kawan* | *sabhangka, bhai* | *sabhangka, bhai* |
| 146 | slave | *budak, hamba* | *bhatua* | *ghata* |
| 147 | dowry | *emas kawin* | *hodea, adhati bhalano* | *adhati bhalano, sara-sara* |
| 148 | rope | *tali* | *habuta* | *rabuta* |
| 149 | road | *jalanan* | *kaangka'a* | *kaangkaha* |
| 150 | boat, canoe | *perahu* | *bhangka* | *bhangka* |
| 151 | knife | *pisau* | *piso* | *piso* |
| 152 | machete | *parang* | *kapulu* | *kapulu* |
| 153 | mortar | *lesung* | *katumbu* | *katumbu* |
| 154 | to pound | *menumbuk* | *de-tumbu* | *de-tumbu* |
| 155 | salt | *garam* | *gaha* | *ghohia* |
| 156 | sugar | *gula* | *gola* | *gola* |
| 157 | what? | *apa* | *o 'ae, o ae* | *o hae* |
| 158 | who? | *siapa* | *la'ae* | *lahae* |
| 159 | where? | *di mana* | *ne 'amai* | *ne hamai* |
| 160 | how? | *bagaimana* | *peda 'ae, nengke 'ae* | *peda hae* |
| 161 | why? | *mengapa* | *noafa* | *noafa* |
| 162 | how much? | *berapa* | *se'ae, fee** | *sehae* |
| 163 | to repeat | *mengulang* | *de-fendua* | *de-fendua* |
| 164 | thirsty | *haus* | *no-kele wu'u* | *do-aha* |
| 165 | to drink | *minum* | *do-foho'u* | *do-foroghu* |
| 166 | hungry | *lapar* | *do-'aho* | *do-gharo* |
| 167 | to eat | *makan* | *do-humaa* | *do-fumaa* |
| 168 | to bite | *menggigit* | *de-sia* | *de-sia* |
| 169 | to see | *melihat* | *de-woha* | *de-wora* |
| 170 | to hear | *mendengar* | *do-fetingke* | *do-fetingke* |
| 171 | to know | *tahu* | *do-pande-'a-ane* | *do-pande-ha-ane* |
| 172 | to sleep | *tidur* | *do-odo* | *do-lodo* |
| 173 | to lie down | *berbaring* | *de-ndole-ndole* | *de-ndole-ndole* |

| | English | Indonesian | Muna – south | Muna – north |
|---|---|---|---|---|
| 174 | sleepy | *mengantuk* | *de-mpau* | *de-mpau, no-tuturu mata* |
| 175 | to dream | *bermimpi* | *do-monifi* | *do-monifi* |
| 176 | to get up | *bangun* | *do-wanu* | *do-wanu* |
| 177 | to kill | *membunuh* | *de-pongko, de-feka-mate* | *de-pongko, de-feka-mate* |
| 178 | to die, dead | *mati* | *do-mate* | *do-mate* |
| 179 | to swim | *berenang* | *do-leni* | *do-leni* |
| 180 | to fly | *terbang* | *do-lola* | *do-horo* |
| 181 | to go | *pergi* | *do-kala* | *do-kala* |
| 182 | to arrive | *datang* | *do-hato* | *do-rato* |
| 183 | to sit | *duduk* | *de-ngkoha* | *de-ngkora* |
| 184 | to stand | *berdiri* | *do-ehe* | *do-ere* |
| 185 | to cough | *batuk* | *do-kangkese* | *do-hoda* |
| 186 | to spit | *meludah* | *de-kapeha* | *do-foghoniu, de-puntori** |
| 187 | to vomit | *muntah* | *do-tongka* | *do-tongka* |
| 188 | to speak | *berbicara* | *do-pugau, do-bisaha* | *do-pogau, do-bisara* |
| 189 | to sing | *menyanyi* | *de-lagu* | *de-lagu* |
| 190 | to laugh | *tertawa* | *do-futaa* | *do-futaa* |
| 191 | to cry | *menangis* | *do-'ae* | *do-ghae* |
| 192 | to take a bath | *mandi* | *de-kadiu* | *de-kadiu* |
| 193 | to fall | *jatuh* | *do-ndawu* | *do-ndawu* |
| 194 | to itch | *gatal* | *no-koito* | *no-koito, no-moito* |
| 195 | pregnant | *mengandung, hamil* | *no-ko-handa, no-fowoowa* | *no-bhala taghi, no-koniwoowa* |
| 196 | to give | *memberi* | *do-wa-'ao* | *do-waa-ghoo* |

**Notes:**

12 *Kahoo/karoo* is the most commonly used type of betel plant; *gili* refers to a type of betel plant which has only leaves and no fruit.

14 The north has no generic term for bamboo; *koo* and *patu* are the most commonly used; other types include *tari*, *tombula* and *wulu*.

19 *Ghoti* also means 'food' in general; it means only 'food' in the south.

24 The short forms without *-omu/-imu* have dual reference; the long forms have plural reference.

37 *Kahoku/karuku* means 'bush, woods, shrubland'; *kamontu'a/katugha* means 'forest, jungle' and *ponue* (not known in the north) 'virgin forest'.

40 The prefix *de-* is the third person realis subject agreement marker for class *ae*-verbs. In this list dynamic verbs are given in the third person plural; stative verbs (such as 43) are given in the third person singular.

48 *Ka'uhi lumola* is literally 'flying animal'.

63–71 The first form of each numeral is used in counting, the second form is used preceding classifiers and measure nouns.

101–102 *Timbu* and *bhaha* are absolute points on the compass and wind directions; *mata'oleo* and *kansoopa* are used only as relative orientation points, as in *kansoopa-no lambu-ku* 'west of my house'. In the north *timbu* and *bhara* refer to the east and west monsoon respectively.

127 *Totolea* is used for men and boys, *'o'oha* for women and girls.

162 As noted in §4.4, *se'ae* is used as a predicate, *fee* as a modifier preceding classifiers and measure nouns.

186 *Foghoniu* refers to soundless vertical spitting, for example through the cracks in a bamboo floor when sitting in a house. *Puntori* is both used of normal spitting (as when walking), but more commonly refers to healing spitting, that is, blowing and spitting lightly over a sick person or body part as a traditional cure.

Of the 196 items, 20 are non-cognates. There are an additional 12 items in which only one of two given equivalents is cognate and the other is not (e.g. 44, 81, 101, 102). I have counted these as half-cognate, so that the total number of non-cognates is 26. This means the dialects share a cognate percentage of 87%.

## 6 Text

The following text was written by Lukas Atakasi, a native speaker of the southern dialect who was born in 1956 in Lolibu and moved to Lakapera in the early seventies. Notice that this text writes initial glottals, while in the standardised southern orthography initial glottals are left unwritten.

**1** *Ne 'umuhu-ku fitu ta'u nandoo se-honda kadhadhia weo dadi-ku.*
loc age-my seven year be one-CLAS event in life-my

**2** *Wakutuu 'aitu 'a-sikola-mo, 'a-hato-mo ne kalasi dua.*
time that 1sR-school-PERF 1sR-arrive-PERF loc class two

**3** *Se-wakutuu, sa-suli-mani ta-mai-'ao we sikola, ta-ta-buko*
one-time WHEN-return-our(excl) 1peR-come-IO loc school JUST-1peR-eat
*'oti-mani maka ta-kala ta-po-kalalambu.*
food-our(excl) then 1peR-go 1peR-REC-play

**4** *Ka-nea-mani ta-po-bhaguli we toombata-no fokoama'ao-ku.*
NOM-usual-our(excl) 1peR-REC-marble loc yard-POS uncle-my

**5** *No-mpona ta-po-kalalambu no-ke-kele-mo wu'u-ku.*
3sR-long 1peR-REC-play 3sR-RED-dry-PERF throat-my

**6** *'A-pugau-mo ne sabhangka-'i-ku, ''A-[m]oni kadeki 'ao-ho'u'.*
1sR-speak-PERF loc friend-PLUR-my 1sI-go.up first 1sI-drink

**7** *Gaha'a do-fetingke pugau-ku, see-see-mie no-ke-kele dua wu'u-no.*
SURPR 3pR-hear speak-my RED-one-person 3sR-RED-dry also throat-his

**8** *Do-pugau peda 'aitu, kansuhu do-po-'amba-'amba te pulangku.*
3pR-speak like that at.once 3pR-REC-RED-chase loc staircase

**9** *Idi 'a-hato-mo te pasaki.*
I 1sR-arrive-PERF loc high.threshold

**10** *'Olotano namu-namu 'a'e-ku do-u'u-e mai-'ao we panda.*
between RED-thought foot-my 3pR-carry.on.head-it come-IO loc under

**11** *So putaa-'a-no 'a-d[um]oli, kansuhu no-sipulu lima-ku.*
for opposite-LOC-its 1sI-turn at.once 3sR-get.loose arm-my

**12** *We wite maka 'a-wula.*
loc earth then 1sR-open.eyes

**13** *'A-hato we wite, 'a-kamposola.*
1sR-arrive loc earth 1sR-faint

**14** *Hato 'a-pande-'ao lalo-ku, do-songko-wi-kanau kalangka*
arrive 1sR-know-INT heart-my 3pR-cover-LOC-me basket

*bhe fotu-ku no-bhehe do-bha-buusi-ane 'oe.*
with head-my 3sR-wet 3sR-INT-sprinkle-it/with water

**15** *Lima-ku se-mbali-a miina-mo 'a-[m]ooli-e-a 'a-sangke-e.*
arm-my one-CLAS-CL not-PERF 1sI-able-it-CL 1sR-lift-it

**16** *'O suana-ku ta-ne-wakule-mo, no-tampu ne langule-no.*
ART left-my JUST-3sR-hang.limp-PERF 3sR-break loc forearm-its

**17** *No-mai ama-ku, no-sangke-kanau-mo se lambu, maka no-kala*
3sR-come father-my 3sR-lift-me-PERF loc house then 3sR-go

*ne-bhasi kamongkula [m]ande-no me-ago-no ka-tampu.*
3sR-call elder clever-A.PART CA-heal-A.PART NOM-break

**18** *No-suli ama-ku no-po-owa bhe kamongkula, bhe ka-intaha-no*
3sR-return father-my 3sR-REC-bring with elder with NOM-hold-his

*hoo-no dana-no.*
leaf-POS tall.grass-its

**19** *Lima-ku me-wakule-no no-woli-wolita-e bhe no-puhu-si-e.*
arm-my CA-hang.limp-A.PART 3sR-RED-turn.round-it with 3sR-massage-REP-it

**20** *Ka-'osa-no ka-lea-no sampe 'a-fendua toha 'a-kamposola.*
NOM-strong-POS NOM-painful-its so.that 1sR-repeat again 1sR-faint

**21** *'A-fo-wula mata-ku lima-ku pada-mo do-sohami-ane*
1sR-CAUS-open.eyes eye-my arm-my already-PERF 3pR-splint-it/with

*sau maka do-gantu-ane dana.*
wood then 3pR-tie-it/with tall.grass

**22** *Kamongkula sa'itu ta-no-kambo-kamboi, amba-no, 'Pe nao-afa-a,*
elder that JUST-3sR-RED-smile word-his FUT.not 3sI-do.what-CL

*sumanomo se-'olota-no 'ahadhi 'aini kadeki me-tende-tende bhaha'i*
if.only one-between-POS week this first.not IMP-RED-turn or

*me-punda-punda-ane-a.*
IMP-RED-jump-it/with-CL

**23** *Ta-'ome-kala-kala nae-mbali sumanomo me-habu-ane ka-temba.'*
JUST-2sI-RED-go 3sI-can if.only IMP-make-for.it NOM-carry.in.sling

**24** *'Olota-no se-'ahadhi maitu haa-paku se-'oleo 'a-kala 'a-owa*
between-POS one-week that two-time one-day 1sR-go 1sR-bring

*lima-ku se lambu-no kamongkula 'aitu.*
arm-my loc house-POS elder that

**25** *'A-hato giagi-a ta-no-fo-diu-diu-mo wiwi-no, miina*
1sR-arrive always-CL JUST-3sR-CAUS-RED-move-PERF lip-his not
*'a-[m]ande-'a-ane-a bhaha'i ne-basa 'ae.*
1sI-know-INT-it-CL whether 3sR-read what

**26** *Pada 'aitu no-punto-hi lima-ku maka no-'awo-'awoloi-'e.*
after that 3sR-blow-LOC arm-my then 3sR-RED-stroke-it

**27** *Peda kaowu 'amaitu, tamaka dhulu 'oleo dhulu ka-taa namisi-ku.*
like only that but the.more day the.more NOM-good feeling-my

**28** *No-pooli ne-efitu ka-ma'oleo-no, ka-gantu-n lima-ku*
3sR-obtain 3sR-seven.days NOM-afternoon-its NOM-tie-POS arm-my
*no-ali-e-mo.*
3sR-remove-it-PERF

**29** *No-po-owa no-tofa-tofa 'owea-ku no-pugau-mo amba-no,*
3sR-REC-bring 3sR-RED-slap shoulder-my 3sR-speak-PERF word-his
*'No-'uhi-mo itu-a.*
3sR-healed-PERF that-CL

**30** *Nae-mbali-mo toha 'o-po-bhaguli-ane-a.'*
3sI-can-PERF again 2sR-REC-marble-it-CL

**31** *'A-suli se lambu bhe 'a-tende.*
1sR-return loc house with 1sR-run

**32** *No-'ia sepali'a lalo-ku, lima-ku no-suli no-po-tudu.*
3sR-glad very heart-my arm-my 3sR-return 3sR-REC-link

**33** *'Ampa 'aitu kamongkula 'ana'a no-mate-mo, tamaka ka-taa-no*
until now elder that 3sR-die-PERF but NOM-good-POS
*lalo-no sa-daa-daa 'a-fe'ulai-'e.*
heart-his ALWAYS-RED-continous 1sR-remember-it

## Free translation

**1** When I was seven years old something happened in my life. **2** At that time I was already going to school, I was in grade two. **3** One time, as soon as we got home from school, we just had a quick bite (lit: we just ate our food) and then went to play. **4** We usually played marbles in my uncle's yard.

**5** When we had played for a long time, I got thirsty. **6** I said to my friends, 'I want to go up into the house to have a drink'. **7** When they heard me say this, each one of them was also thirsty. **8** They spoke like that and at once they all ran for the staircase (leading into the house). **9** I had already arrived at the high threshold board. **10** Then unexpectedly my legs were pushed up from below (lit: carried on the head from below). **11** The moment I wanted to turn round, my hand slipped off. **12** It was on the ground that I opened my eyes. **13** On the ground I passed out. **14** When I regained consciousness, they covered me with a basket and my head was wet being sprinkled with water. **15** I could not lift one of my arms. **16** My right hand was just hanging limp, it was broken at the forearm.

**17** My father came and carried me home, then he called an old man who knew how to heal fractures. **18** My father returned together with the old man, who held leaves of tall grass in his hand. **19** He turned my limp arm around repeatedly and massaged it. **20** It hurt so much that I passed out again. **21** When I opened my eyes, my arm had been splinted with a piece of wood and tied together with tall grass. **22** The old man just smiled and said, 'It'll be all right, but this coming week you mustn't run or jump with it. **23** Just walking is fine, as long as you make a sling for it'. **24** Twice a day that week I went to the old man's house to have my arm treated (lit: to bring my arm). **25** When I arrived, he would always only move his lips, I didn't know what he was saying. **26** After that he would blow over my arm and stroke it. **27** That was all, but day by day I felt better.

**28** In the afternoon of the seventh day he removed the splint from my arm. **29** While he slapped my shoulder he said, 'It is healed. **30** You can play marbles again with it'. **31** I ran back home. **32** I was so happy that my arm was in one piece again.

**33** That old man is already dead, but I continue to remember his goodness.

**Notes:**

**14**. When someone faints from falling, it is customary to put a basket over that person and sprinkle water over him or her, to ward off further harm.

**25**. 'What he was saying', literally 'what he was reading'; the verb *basa* 'read' is used because the old man was muttering traditional charms and prayers. The full term is *debasa dho'a* 'to recite a charm'.

## Abbreviations

| | | | |
|---|---|---|---|
| A.PART | active participle | loc | locative preposition |
| ART | article | LOC | locative |
| BI | Bahasa Indonesia | N | northern dialect |
| CA | class affix | NOM | nominaliser |
| CAUS | causative | p | plural (inclusive) |
| CL | clitic | PERF | perfective |
| CLAS | classifier | PLUR | plural |
| DETR | detransitiviser | pol | polite |
| du | dual | POS | possessive linker |
| DU | Dutch | P.PART | passive participle |
| e, excl | exclusive | PURP | purpose |
| EMPH | emphatic particle | R | realis |
| FUT | future | REC | reciprocal |
| I | irrealis | RED | reduplication |
| IMP | imperative | REP | repetitive |
| incl | inclusive | s | singular |
| INT | intensifier | S | southern dialect |
| IO | indirect object | SURPR | surprise particle |
| lit | literally | Wol | Wolio |

## Bibliography

Anceaux, J.C., 1952, *The Wolio language: outline of grammatical description and texts.* The Hague: Nijhoff.

—— 1987, *Wolio dictonary (Wolio–English–Indonesian). Kamus Bahasa Wolio (Wolio–Inggeris–Indonesia).* Dordrecht: Foris.

Berg, René van den, 1989, *A grammar of the Muna language.* Dordrecht: Foris.

—— 1991a, Muna dialects and the Munic languages: towards a reconstruction. In Ray Harlow, ed. *VICAL 2. Western Austronesian and contact languages: papers from the Fifth International Conference on Austronesian Linguistics*, 21–51. Auckland: Linguistic Society of New Zealand.

—— 1991b, Muna historical phonology. In J.N. Sneddon, ed. *Studies in Sulawesi linguistics,* part 2, 1–28. Jakarta: Badan Penyelenggara Seri Nusa, Universitas Katolik Indonesia Atma Jaya. *NUSA* volume 33.

Berg, René van den, in collaboration with La Ode Sidu, 1996, *Muna–English dictionary.* Leiden, KITLV Press.

Couvreur, J., 1935, Ethnografisch Overzicht van Moena. Unpublished typescript.

—— 2001, *Sejarah dan Kebudayaan Kerajaan Muna.* Kupang, Artha Wacana Press.

Petyt, K.M., 1980, *The study of dialect.* London: Deutsch.

Vonk, H.W.,1937, Nota betreffende het Zelfbesturend landschap Boeton. Unpublished typescript.

Van {den, R.}. "Notes on the southern Muna dialect". In Bowden, J. and Himmelmann, N. editors, *Papers in Austronesian subgrouping and dialectology*.
**PL-563:129-169**. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-563.129*